Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 2744706, 29 pages
https://doi.org/10.1155/2023/2744706

Research Article

Localization and Detection of Multiple Attacks in Wireless Sensor Networks Using Artificial Neural Network

Gebrekiros Gebreyesus Gebremariam,[1,2] **J. Panda**,[2] **and S. Indu**[2]

[1]*Department of Electronics and Communication Engineering, Raya University, Maychew, 7020 Tigray, Ethiopia*
[2]*Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Main Bwana Road, Delhi 110042, India*

Correspondence should be addressed to Gebrekiros Gebreyesus Gebremariam; kiros2004comp@gmail.com

Received 11 June 2022; Revised 30 November 2022; Accepted 19 December 2022; Published 10 January 2023

Academic Editor: Javier Prieto

Security enhancement in wireless sensor networks (WSNs) is significant in different applications. The advancement of routing attack localization is a crucial security research scenario. Various routing attacks degrade the network performance by injecting malicious nodes into wireless sensor networks. Sybil attacks are the most prominent ones generating false nodes similar to the station node. This paper proposed detection and localization against multiple attacks using security localization based on an optimized multilayer perceptron artificial neural network (MLPANN). The proposed scheme has two major part localization techniques and machine learning techniques for detection and localization WSN DoS attacks. The proposed system is implemented using MATLAB simulation and processed with the IBM SPSS toolbox and Python. The dataset is classified into training and testing using the multilayer perceptron artificial neural network to detect ten classes of attacks, including denial-of-service (DoS) attacks. Using the UNSW-NB, WSN-DS, NSL-KDD, and CICIDS2018 benchmark datasets, the results reveal that the suggested system improved with an average detection accuracy of 100%, 99.65%, 98.95%, and 99.83% for various DoS attacks. In terms of localization precision, recall, accuracy, and f-score, the suggested system outperforms state-of-the-art alternatives. Finally, simulations are done to assess how well the suggested method for detecting and localizing harmful nodes performs in terms of security. This method provides a close approximation of the unknown node position with low localization error. The simulation findings show that the proposed system is effective for the detection and secure localization of malicious attacks for scalable and hierarchically distributed wireless sensor networks. This achieved a maximum localization error of 0.49% and average localization accuracy of 99.51% using a secure and scalable design and planning approach.

1. Introduction

Hierarchically distributed microsensor nodes in the field are linked together with multihop wireless communication technologies to form a wireless sensor networks (WSNs) [1]. The sensor nodes have sensing and wireless communication modules, storage, and processing units. Wireless sensor networks (WSNs) monitor and collect network information, including network state and data transmission. They also monitor object positioning and tracking [2]. WSNs are vulnerable to multiple, including Sybil attacks, wormhole attacks and eavesdropping [3]. The Sybil attack is the most harmful routing attack that fabricates and depreciates multiple fake identities launching a malicious attack on the legitimate node to lower service quality [4, 5]. It is the key research challenge in wireless sensor networks like the other applications, including architecture, healthcare, disaster management, deployment quality of service, calibration [6], and synchronization. Wireless sensor nodes sense and record data from the environment and send the data to the cluster head for the aggregation process. An intelligent sensing and computing framework is essential for the security localization and detection of attacks using an artificial neural network (ANN). This scheme is gaining attention due to its low computational cost and faster convergence. In this paper we proposed an optimized multilayer perceptron artificial

neural network for detection and localization of routing attacks in WSNs.

Emerging applications of wireless sensor networks (WSNs) include traffic management and object tracking, both of which require localization of sensor nodes [7]. For efficient routing and location-aware services, it is crucial to have an accurate estimation of the sensor node location. Without knowing where the sensor is located, the data collected by WSN is often useless. As a result of their potential utility in a wide range of WSN applications, the localization techniques are garnering a rising amount of attention from researchers. Based on the information required, localization techniques can be divided into two broad categories: range-based techniques, which rely on the known distances or angles between nodes to make location estimates, and range-free techniques, which instead make location estimates based on the closeness of a number of reference nodes. Range-free approaches are replacing range-based methods in WSN localization due to their lower hardware and computational requirements. The clustering and localization algorithm is a well-known example of a range-free technique since it uses adjacent reference nodes' positions to estimate the node's own location. During the setup phase, the starting positions of the reference nodes are either hard-coded or computed.

1.1. MLPANN Applications in WSNs. Network traffic identification is becoming increasingly popular as a field of study in the field of network administration, drawing researchers from all over the world [8]. The growth in network capacity is closely proportional to this rise. New forms of network applications, such as peer-to-peer (P2P) file sharing, have rendered some of the more traditional methods of traffic identification, including port-based or deep packet inspection, ineffective. Selecting an appropriate feature selection method, which can select the best features according to the impact of the great traffic behaviour characteristics, is essential for achieving higher recognition efficiency and greater identification accuracy when performing traffic identification based on ML (Machine Learning). The MLP (Multilayer Perceptron) approach is superior to other identification algorithms in terms of identification accuracy. It has been shown that as the number of training samples grows, the identification rate also rises. But MLP is not without its own value and drawbacks that mean it needs to be enhanced.

There is now a great deal of academic focus on creating fingerprint localization algorithms that make use of artificial neural networks (ANNs). Despite noisy RSSI measurements, the ANN can still give accurate recognition of the node's position, which is a major advantage. Using ANNs eliminate the need for detailed knowledge of the indoor environment or the locations of reference nodes. In order to approximate a mapping between the multidimensional fingerprint space and the coordinates of nodes, ANN interpolates the acquired data in the fingerprint database. During the ANN's training process, the collected RSSI vectors are used to fine-tune the weights of connections between neurons. Although training may take a while, the localization process is far faster than any analytical estimates of the node's position.

Multilayer perceptron (MLP) is the most widely utilized ANN architecture in modern range-free wireless sensor node localization applications. In WSN, fingerprint-based localization was accomplished using the MLPANN. For this evaluation, we evaluated 43 distinct backpropagation training algorithms to determine the method's accuracy. Something very close to that was proposed as a tactic. The ANN training has been kept up-to-date at regular intervals so it can adjust to changing conditions on the wireless channel. Four MLPANNs, each with a different amount of inputs, formed an ensemble that was shown to the audience. This tactic specifies that, if the localization operation must be executed, an ANN with the same number of inputs as the currently connected reference nodes is selected and issued. Given the poor scalability of this approach, we settled on a cap of four connections between reference nodes. When compared to methods based on fuzzy learning systems or genetic algorithms, the localization results generated by the ANNs ensemble were shown to be more reliable. Here we propose a cooperatively optimized and secure multilayer perceptron artificial neural network (MLPANN) for scalable and broad area networks that combines range-based and range-free localization technique approaches from the realm of artificial intelligence (AI) to problems inherent to wireless sensor networks (WSNs) [9], such as data aggregation and fusion, routing, task scheduling, optimal deployment, and localization as shown in Figure 1. In this context, the term "computational intelligence" refers to a subfield of machine learning that combines techniques with roots in biology, such as neural networks, fuzzy systems, and evolutionary algorithms, to develop forecasting models. This algorithm for learning could be built using cascading decision chains for recognizing nonlinear and complex functions. However, the high-computational requirements for learning the network weights and the substantial administrative overhead mean that distributed neural networks are not yet widely used in WSNs. Neural networks, on the other hand, are well-suited for handling many network difficulties with a single model because of their ability to simultaneously learn multiple outputs and decision boundaries in centralized solutions.

1.2. WSN Routing Attacks. Attacks on the wireless sensor network's network layer can hinder performance by rendering a genuine node unreachable to the service. The following section will go through some of the more common types of attacks seen in WSNs. With these kinds of attacks, sensitive data is compromised before it even reaches the target node. The attacks in wireless sensor networks operate in all layers of the network exploiting resources and degrading the service quality of the network.

1.2.1. Wormhole Attack. The wormhole attack is created by two malicious nodes having a tunnel path in the two locations and misapprehension. Wormhole attack attracts and manipulates significant network data traffic launching various attacks. It advertises its packets using the intermediate nodes to sniff, modify, and drop from reaching the destination [10]. Figure 2 cluster A shows a wormhole attack scenario depicting when source node S sends the packet to the sink node. A

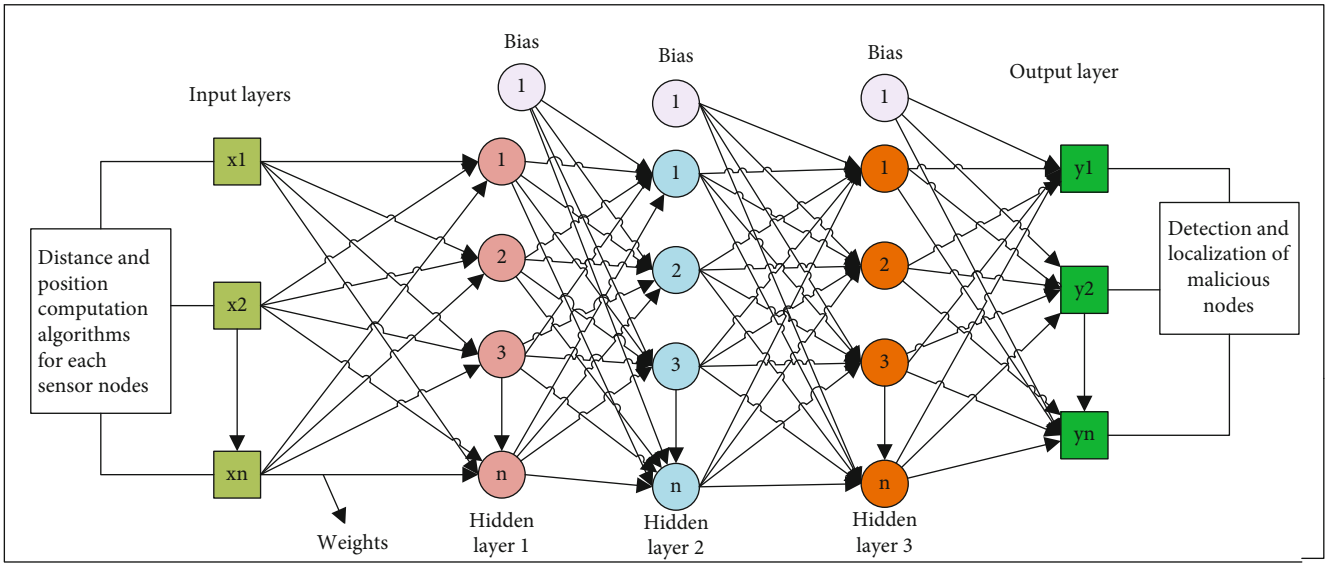


Figure 1: Illustration of MLPANN for detection and localization of WSN attacks based on localization algorithms.

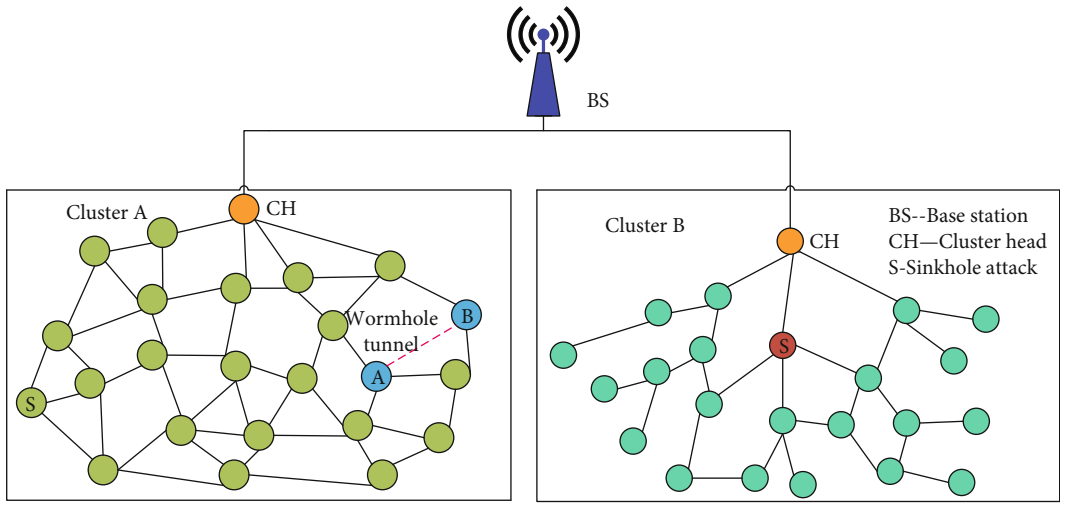


Figure 2: Illustration of wormhole and sinkhole attacks in WSNs in cluster A and B, respectively.

wormhole tunnel is created between malicious nodes A and B. The packet is dropped and modified by the tunnel before reaching the base station. At least the two hostile nodes employing a secure communication channel known as a tunnel can detect a wormhole assault [11]. It is at this point that the wormhole tunnel will begin to collect the data packets and forward them on. The malicious node on the other end of the tunnel receives a control packet. At the other end, it uses a private channel to relay the packet to another node that has caught its attention. For enhanced metrics, such as fewer hops or less time, the private channel is selected as the conduit for communication between the source and the destination. The attack usually consists of two phases. Multiple initial directions are of relevance to the wormhole nodes. The second stage is when the packets begin to make use of the malicious nodes. It is possible for these nodes to hinder the network's performance in a number of ways. Wormhole nodes can be used to steal information or communicate it to a third party if they delete, tamper with, or send it.

1.2.2. Sinkhole Attack. Sinkhole attack advertises routing paths to the base stations, making itself a normal node misguiding the neighbor nodes that cause threats to the network. The malicious nodes create a hole in the routing path that can damage the regular operations of the network. The sinkhole attack uses a compromised node with fewer hops to advertise the route to the destination. This misuse of routing information misguides the legitimate node and attracts the node closer. Figure 2 cluster B illustrates the scenario of a sinkhole attack for attracting and capturing packets from the neighbor nodes. The sinkhole attack utilizes a secret tunnel for attracting nodes and capturing packets. The malicious node then deceived and sent packets to the base station.

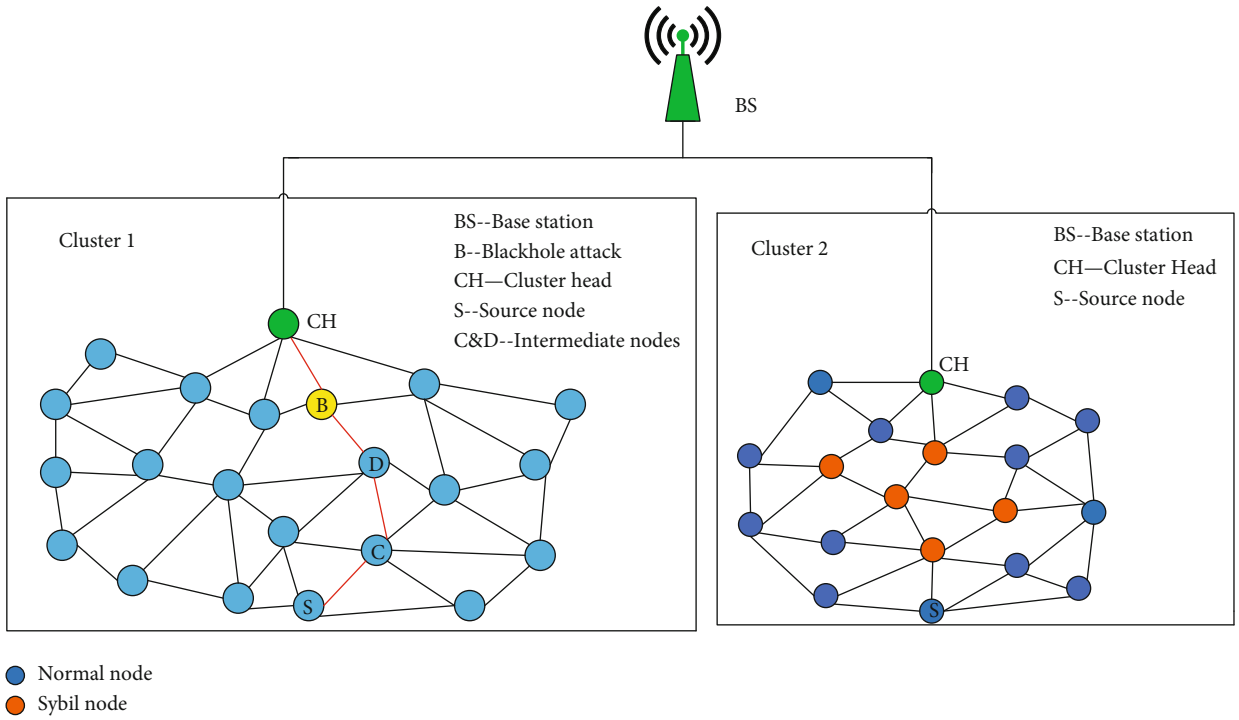


Figure 3: Illustration of blackhole and Sybil attack scenario in WSNs for cluster 1 and cluster 2, respectively.

1.2.3. Blackhole Attack. Blackhole attacks capture and reprogram sensor nodes to block packets instead of receiving and forward to the base station [10]. The blackhole attack compromises the information with the malicious node that enters the blackhole region. The blackhole attack undermines the network performance by using the network partitioning so the essential updates cannot reach the base station. It degrades the network performance metrics and consumes large network traffic. The source node S sends a packet to the base station using the intermediate C and D nodes, as shown in Figure 3 of cluster 1. The blackhole node consumes the entire traffic and is not forwarded to the destination node.

The blackhole attack performs suspicious activity using loopholes for discovering the routing [12]. The blackhole attack compromises the legitimate node with a malicious node so that the packets are dropped and unable to reach the destination nodes. The suspicious node cause packet is dropping for targeted nodes and customizing set of nodes for packet dropping. The information that comes to the blackhole is dropped and sent fake packets to the base station. The routing requests and routing response messages by the blackhole attacks have higher order number which is greater than the normal node request and response so that the normal node will not respond to the routing request with higher order number which causes deletion of routine from the networks [13].

1.2.4. Sybil Attack. Sybil attack forges and spoofs the identity of the legitimate node in wireless sensor networks [14]. Sybil attack interrupts the routing table and the trust value of the node of the legitimate node. Sybil attack duplicates multiple identities for confusing the neighbor nodes [15]. This attack uses geographic routing protocols for targeting authorized nodes. Sybil attack takes numerous identities to disguise the storage entities of the legitimate node [16], as shown in Figure 3 of cluster 2. The malicious node transmits data with imaginary events multiple times. This type of attack creates illusion that makes it difficult to detect the whole network.

1.3. Problem Formulation. Most of the existing literature review papers dealt with single attacks with low localization and detection accuracy in WSNs. Thus the deployment of wireless sensor nodes seeks optimal and intelligent localization methods for accurate node position and attack identification. To overcome this problem, it is essential to design and implement a new effective technique. This paper proposes a security localization and detection scheme employing optimized multilayer perceptron artificial neural network for various classes of attacks against wireless sensor networks, which are vulnerable to a wide variety of denial-of-service (DoS) attacks that exploit the network's resources. The proposed method is designed for multiple attack detection and classification, representing the input and output relationships using the ANN technique. The design and planning of distributed hierarchical clustered topology are also discussed that consists of sink node, cluster head, malicious and sensor nodes as shown in Figure 4. Also, we discuss the effectiveness of the proposed system using the benchmark datasets including CICIDS2018, UNSW-NB 15, WSN-DS and NSL-KDD with different evaluation metrics using training and testing samples as a benchmark for performance evaluation. The dataset is processed using batch mode.

1.4. Research Contribution. The proposed attack localization and detection scheme is based on an optimization multilayer

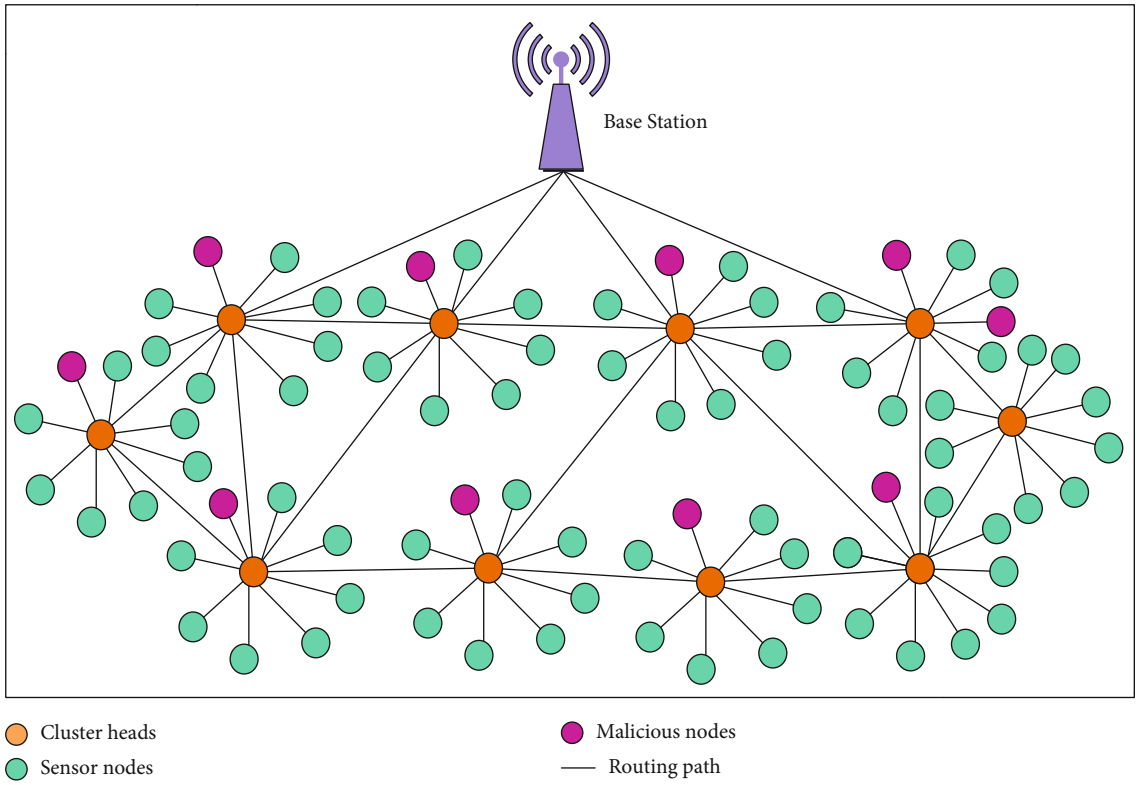


Figure 4: Clustering in hierarchically distributed wireless sensor network model.

perceptron neural network [17]. The proposed system possesses different phases with the proper network planning and node configuration. These include network data processing and feature extraction, training, and testing for attack detection and classification. Some of the novel contributions of this work are as follows:

(1) To design and simulate wireless sensor network topology with attack detection and localization features

(2) To explore the various routing attacks and techniques simulating these attacks using clustering and routing protocols

(3) Evaluate the network performance using a sample public dataset as a benchmark with attack detection localization metrics

(4) Explore machine learning techniques secure localization and detection of routing in wireless sensor networks in all layers of the network

(5) A multilayer perceptron neural network technique enables the detection and classification of malicious nodes using network traffic data and feature extraction. It maximizes the location and position accuracy of the suspicious node

(6) To detect and localize multiple attacks with greater classification accuracy for clustered and hierarchical network architecture

(7) Measure the security performance of the scheme using comparison performance for effective validation and confirmation with similar previous works

(8) Explore hybrid range-based and range-free localization techniques for unknown and malicious nodes that affect quality of service in WSN using collaborative approach

The rest contents of this paper are organized into different sections and structures. Section 2 encompasses the previous literature works. Section 3 describes the network and attack models depicting graphically in detail using clustering and routing protocols. The next part is Section 4, which discusses the proposed attack localization and detection technique in WSNs using MLPANN approach. The next Section 5 details the simulation and experimental analysis using a benchmark dataset for different classes of routing attacks. The last section is Conclusion and remarking for future works.

2. Related Works

Messous and Liouane [1] presented an online successive distance vector hop scheme for node localization accuracy in WSNs. They also discussed the variation of anchor nodes with optimized distance between nodes in the network. Dong et al. [2] examined the distance vector hop algorithm against Sybil attacks for effective node localization and

accuracy for improved security in WSN. The scheme also reduces the average error localization by 3%, setting the beacon nodes 50 in the simulation that is 78%. Chelouah et al. [18] addressed localization algorithm in mobile WSNs. They also presented the mobility of nodes for coverage optimization, connectivity, and analysis. Hadir et al. [19] presented a localization technique in WSNs using an effective distance vector hop scheme. They also discuss the average hop size and localization accuracy by exploiting the information. Almomani et al. [20] designed a low cost and efficient, intelligent DoS attack detection and prevention technique. They also discuss different DoS attack classifications using a specialized dataset for WSN. Patel and Mistry [21] presented Sybil node detection [22] using various schemes. They also discussed and analyzed the protocols used in WSNs. Yavuz et al. [23] proposed detecting IoT-routing attacks using a deep learning machine learning technique. The Cooja simulator generates high-fidelity attack data in the IoT network with 1000 sensors. Sujatha and Anita [24] examined the detection of Sybil attack detection using hybrid fuzzy and powerful extreme learning machines. They also discussed ARM as the main CPU with LEACH environment and ZigBee transceivers on real-time testbeds. Qi et al. [25] researched a localization algorithm to improve the node position accuracy and reducing localization error in WSNs using MA-MDS. They also use the Prussian analysis algorithm for accurate coordinate transformation. Li et al. [26] presented a localization trust valuation scheme to detect spoofing and Sybil attacks. This scheme is obtained by selecting localization performance, estimated distance, and transmission with the threshold property set in WSNs. Song et al. [27] proposed a chaotic hybrid mutation and chaotic inertial weight-updating technique with a glowworm swarm optimization approach. The scheme also avoids premature convergence with better convergence and higher accuracy. Saud Khan and Khan [28] presented Sybil attack detection using signed response authentication techniques for global mobile communication systems. They also discussed the probabilistic model to analyze Sybil attack detection performance.

Abbreviations, acronyms, and shorter variants of terms and phrases used as resources in this paper are included as shown in Table 1

3. Network Model

Sink, cluster head, sensor, and attacker nodes are all represented in the simulated network. The sensor nodes tend to group together throughout the system. When sending information to the beacon nodes and the base station, each cluster chooses its own cluster head node to act as the central hub for that cluster. The beacon nodes decide on the optimal routing path by employing an optimization strategy based on a fitness function. In this context, the terms beacon nodes and anchor nodes are synonymous. Sybil attacks and wormhole assaults are the types of attacks that can be used against this network paradigm. The positioning and placement of anchor nodes are dependent on their relationship to other nodes. Once an anchor node has been put in a network, it will remain in the same location permanently. The sensor nodes have claimed their territory. The localization scheme is used for providing accurate position and location of the sensor nodes by making clustering of the nodes having one cluster head for each group [29, 30]. Unknown nodes locate their positions using the anchor node assistance. Sensor nodes update their locations periodically by the system. Malicious nodes broadcast their positions by creating multiple fake identities and advertising themselves as the beacon nodes. The malicious node also creates tunnels for dropping packets before reaching the destination. The hierarchical clustering of the sensor deployments enables less energy consumption and enhances the network life time as shown in Figure 4.

The legitimate nodes are assumed to be homogenous in computational processing, storage capacity, communication level, and activation energy in the model [31]. Malicious nodes are considered more effective than the legitimate node for the activity in capturing the security key of the base station and clusters. The attacker disrupts the normal functioning of the network by cloning the authorized node.

3.1. Cluster Formation and Data Aggregation. Clustering is a method of organizing a set of sensors to increase the durability of the network and decrease its power consumption [32]. The network's sensor nodes are organized into groups of similar devices. Collectively, the sensor nodes that make up the cluster gather data and send it on to the cluster coordinator. The data is aggregated and filtered by the cluster head before being sent to the hub. The sensor's stable functioning and neighbor evaluation are aided by the clustering method. Cluster heads (CHs) are the most important nodes in the cluster since they serve as the hub for monitoring. The three selection criteria are used to determine which of the sensor nodes will become the cluster head.

(i) The number of nearby neighbor nodes

(ii) The quality of the received signal from the sensor node

(iii) The node's remaining energy before it is activated

(iv) The minimum distance to the base station as calculated by the distance vector protocol. The distance D between any two sensor nodes is computed using the distance vector technique as shown below

$$D = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}. \quad (1)$$

The u and v coordinate for nodes i and j, respectively. Computing the distance between the base stations to any node with a small distance is likely the cluster head. The energy employed for the communication and activation of the network model is evaluated by setting the threshold parameters with the multipath model approach. The amount of energy E_{TX} for k-bits of data transmission over D distance and D_o threshold distance is given as in

Table 1: List of abbreviations and acronyms used in this paper.

Abbreviations	Descriptions
BS	Base station
WSNs	Wireless sensor networks
CH	Cluster head
D	Distance
E	Energy of the sensor node
RSSI	Received signal strength indicator
DV-H	Distance vector hop
IoT	Internet of Things
ANN-IDS	Artificial neural network-based intrusion detection system
FEMs	Fuzzy extreme machines
MK-ELM	Multikernel extreme learning machine
LEACH-ANN	Low-energy adaptive clustering hierarchy based on ANN
CNN-MCL	Convolutional neural network and mean convolutional layer
HTM-LSTM	Hierarchical temporal memory and long short-term memory
MLPANN	Multilayer perceptron artificial neural network
RL-IDS	Reinforcement learning-based IDS
GBFS-IDS	Gradient boosting feature selection for IDS
ML-ID	Machine learning-based intrusion detection
UWB	Ultrawide band
DI-ADS	Deep intelligent attack detection scheme
PO-CFNN-	Political optimizer based on cascade forward neural network
RNN	Recurrent neural network
ECGAL	Energy-efficient clustering and localization centered on genetic algorithm
LSTM-FFNN	Long short-term memory and feed-forward neural network
DNN-CSO	Deep neural networks with chicken swarm optimization

$$E_{TX} = \begin{cases} k \times E_e + k \times E_f \times D^2, & \text{if } D \le D_o, \\ k \times E_e + k \times E_m \times D^4, & \text{if } D > D_o, \end{cases} \quad (2)$$

where E_{TX} is the transmitted energy, E_f is the reception energy, and E_e is the power dissipated in the transmitter or receiver for single-bit data transmission. The dissipated energy depends on signal spreading, filtering, modulation, and channel-coding factors. The threshold transmission distance D_o with k-length of data transmission is given by

$$D_o = \sqrt{\frac{E_f}{E_m}}. \quad (3)$$

For k-bits of message reception, the energy consumed by the receiver node is

$$E_{\mathrm{RX}(k)} = k \times E_e. \quad (4)$$

3.2. Localization Techniques. Numerous wireless sensor network (WSN) applications rely on localization to locate a target by comparing the signal strengths of transmitters and receivers already set up in the region of interest [33, 34]. Some algorithms are essential for finding and assessing the location and position of the nodes and security enhancement for precise location of the target. The scheme is divided into range-based and range-free localization techniques. The latter one is cost-effective with special hardware requirements. The received signal strength indicator (RSSI) and distance vector hop localization algorithms evaluate wireless sensor node accurate position and location. The distance vector localization procedure is essential to compute the coordinates of the sensor nodes and cluster heads using the beacon nodes [1]. The scheme calculates and manipulates the position and distance of the unidentified nodes. The distance vector hop procedure helps to find the spaces among the beacon nodes in WSN. The calculated minimum distance is the average hop size using the distance vector approach. This algorithm was first identified by [2]. The distance vector localization scheme is a range-free strategy [19] with a series of steps in Table 2.

The average distance hop for the anchor node is computed and obtained relative to another beacon with the minimum hop count given by

$$\mathrm{HS}_i = \frac{\sum_{i \ne j} \sqrt{(ui - uj)^2 + (vi - vj)^2}}{\sum_{i \ne j} hij}. \quad (5)$$

The interpretation of the variables is shown in Table 3.

Table 2: Steps in distance vector technique for computing the position of unknown nodes.

Steps	Description of steps
Routing initialization	Beacon node broadcasts to all nodes and the hop count of its location information with 0 value initialized in the network [19].
Calculating distance	The average size of the hop and the unknown node distance is calculated with the assistance of the beacon node.
Position estimation	Unknown nodes estimate their position by the triangulation algorithm or maximum likelihood estimators or polygon method with the help of the anchor node.

Table 3: Equation (1) variable description and interpretation.

Variable	Description of the variable
i, j	True anchor nodes
(ui, vi), (uj, vj)	The known and true coordinates for i and j
Hij	Hop counts of the anchor nodes
HSi	Average hop distance

Anchor node transmits its information [2] followed by hop-size calculation. The distance between sensor node and anchor computed with hop-size details is given by

$$D_{pk} = HpSi_{i\times} hp_{pk}. \tag{6}$$

The polygon technique enables the estimation of the position (P) of each anonymous node. P is the spot of the unidentified node denoted as (u, v) and di, the space among anchor and indefinite nodes. The position of the strange node p assuming n beacon nodes involved is estimated by [2].

$$\begin{aligned}
(u-u_1)^2 + (v-v)^2 &= {D_1}^2,\\
(u-u)^2 + (v-v)^2 &= {D_2}^2,\\
&\vdots\\
(u-u_n)^2 + (v-v_n)^2 &= {D_n}^2.
\end{aligned} \tag{7}$$

We can get a set $(n-1)$ of expressions subtracting from the first equations to make the system linear, given as depicted in

$$\begin{aligned}
{u_1}^2 + v^2 - {u_n}^2 - {v_n}^2 - 2(u-v_n)u - 2(v_1-v_n)v &= {D_1}^2 - {D_n}^2,\\
{u_2}^2 + {v_2}^2 - {u_n}^2 - {v_n}^2 - 2(u_2-u_n)u - 2(v_2-v_n)v &= {D_2}^2 - {D_n}^2,\\
&\vdots\\
{u_{n-1}}^2 + {v_{n-1}}^2 - {u_n}^2 - {v_n}^2 - 2(u_{n-1}-u_n)u - 2(v_{n-1}-v_n)v &= {D_{n-1}}^2 - {D_n}^2.
\end{aligned} \tag{8}$$

Rearranging the previous equations into the formula of ui $= \mathrm{BA}^{-1}$, where A, ui, and B are expressed as in

$$A = \begin{cases} 2(u-u_n)2(v_1-v_n), \\ 2(u_2-u_n)2(v_2-v_n), \\ \vdots \\ 2(y_{n-1}-u_n)2(v_{n-1}-v_n), \end{cases} \tag{9}$$

$$B = \begin{cases} {u_1}^2 + v^2 - {u_n}^2 - v^2 + {D_n}^2 - {D_1}^2 \\ {u_2}^2 + v^2 - {u_n}^2 - {v_n}^2 + {D_n}^2 - {D_2}^2 \\ \vdots \\ {u_{n-1}}^2 + {v_{n-1}}^2 - {u_n}^2 - {v_n}^2 + {D_n}^2 - {D_{n-1}}^2 \end{cases}, \tag{10}$$

$$U_i = \left(\frac{u}{v}\right). \tag{11}$$

The location of the node is computed solving the least square method stated as in

$$U = \left(A'A\right)^{-1} A'B. \tag{12}$$

The clustering and distance vector-routing protocols are used in the proposed scheme for effective wireless sensor network deployment.

In contrast to its range-based equivalents, RSSI-based localization algorithms have gained a lot of traction in the academic community for a variety of compelling reasons [35]. Today's wireless sensor nodes typically include features like RSSI measurement and data transmission to higher stack layers. For RSSI-based localization, no time-synchronization between nodes, ultrawide band (UWB) radios for more precise time of arrival calculations, or antenna arrays are needed. In terms of both software and hardware, it is a straightforward and inexpensive approach to achieving node localization. However, the DV-hop algorithm completely skips measuring the real distances between the one-hop neighbor nodes and leveraging these distances for more precise localization in massive-scale wireless sensor networks.

To localize wireless nodes using the DV-hop technique, the hybrid approach takes two extra steps, as indicated above for improvement of the localization accuracy and malicious

node detection. Instead of relying on the average hop distance like the original DV-hop algorithm did, we first use the RSSI data to estimate the distances between the anchor nodes and their one-hop surrounding sensor nodes. Using the RSSI value does not necessitate any specialized hardware or additional expenditures because the MAC sublayer in most modern wireless sensor nodes computes RSSI value for every received packet and sends that value to higher layers. Second, after a sensor node N has been located, it is elevated to the role of anchor, which is utilized to localize other sensor nodes. With more (repurposed) anchor nodes to work with, the remaining sensor nodes can be localized with greater precision. That is especially useful in wireless networks when there are fewer anchor nodes. Third, differential evolution (DE) is a technique used in evolutionary computation to find the optimal solution to a problem by iteratively trying to enhance a candidate solution with respect to some quality metric. Metaheuristics are approaches that search enormous spaces of possible solutions while making few or no assumptions about the underlying problem. Unfortunately, metaheuristics like DE cannot promise you will get the best possible result every time.

Since DE does not rely on the gradient of the optimization problem, DE can be applied to optimization problems involving multidimensional real-valued functions even if the problem cannot be differentiated, as is the case with traditional optimization techniques like gradient descent and quasi-newton methods. For this reason, DE can be applied to optimization problems that are inherently noncontinuous, noisy, dynamic, etc. Using its basic equations, DE optimizes a problem by keeping a population of candidate solutions, generating new candidate solutions by merging old ones, and finally keeping the candidate solution with the highest score or fitness on the optimization task at hand. As a result, the gradient is unnecessary because the optimization issue is viewed as a black box that only delivers a measure of quality given a candidate solution.

The problem of the localization of techniques is transformed into multilayer perceptron artificial neural network by computing the distance and position of each type of nodes with unique identity for detection and localization of the malicious nodes as shown in Figure 5. The sensor nodes in our purpose are assumed to both homogenous and heterogeneous wireless sensor networks. The beacon nodes have high-computational data processing and have their own localization that helps for other nodes to estimate and compute their location and positon of the ordinary sensor nodes in the network. Adding machine learning to WSN localization helps increase the precision of range-free node positioning [36]. In particular, the use of artificial neural networks (ANNs) in range-free localization algorithms has significantly improved their accuracy and performance compared to more conventional methods. The MLPANN learning strategy is needed that starts with a labelled dataset in order to construct a model that can appropriately generalize to data that was not included in the training set before we can make any adjustments to the weights [37].

3.3. Attack Model. One hundred nodes are distributed randomly in 1000×1000 m square areas with sensor, cluster head, and malicious nodes. The proposed scheme aims to enhance the detection accuracy of security localization [2] to routing attack using the distance vector hop procedure and clustering protocols in WSN using an artificial neural network approach. Figure 6 is the Sybil attack model with three sets of wireless nodes. A Sybil attack is a type of network assault in which a malicious node purposely and illegally displays a large number of forging or false identities to other sensor nodes [38]. This is accomplished by either independently generating new identities or illegally assuming the identities of other sensor nodes. By creating an unpredictable number of fake node identities, a Sybil node might interfere with WSN operations like multipath routing, which uses a variety of routes to find the best one between a source and a destination. The attack model shows how malicious nodes launch fake behaviors creating multiple identities against the position and location of the legitimate node with various routing paths. This degrades the lifetime of the network by reducing the computational performance of the authorized nodes.

The other routing attacks including scheduling, blackhole, grayhole and flooding attacks are used in for simulating and implementation of the proposed scheme using the WSN-DS dataset as a benchmark for evaluating for localization and detection.

3.4. Benchmark Datasets. In this section, three benchmark datasets including UNSW-NB 15, WSN-DS, and NSL-KDD are utilized to measure the effectiveness attack detection and localization accuracy. The raw network packets of the UNSW-NB 15 dataset were generated by cyber LAB using the IXIA PerfectStorm tool for cyber security for generating attack behaviors [39]. The cyber security dataset [40] is structured into training and testing samples using the batch mode for updating the total error for each weight, as shown below in Table 4. The dataset contains ten classes of attacks with different statistical frequency distribution in the network.

There are various attack activities in the dataset for processing and classification of the proposed system. The attacks are classified as normal, shellcode, analysis, backdoor, backdoors, DoS, exploits, fuzzers, reconnaissance, generic, and worms [39, 42–46].

The frequency distribution of the four types of routing attacks found in the WSN-DS dataset is provided in Table 5 and is utilized as a benchmark against which the performance of the proposed can be measured. There are a total of 84556 data points and 23 features available to use in creating a predictive model, eighty percent are utilized for training, while twenty percent are used for testing.

The other benchmark dataset for evaluating the proposed technique is the NSL-KDD containing 100069 samples and with classes of attacks including denial-of-service (DoS), probes, user to root (U2R), root to local (R2L), and normal as shown in Table 5. The dataset has 41 features with 38 numerical and 3 categorical features.

4. Proposed System

The proposed system consists of a series of phases: design and planning, deployment and routing, data processing, training and testing, attack classification, attack detection,

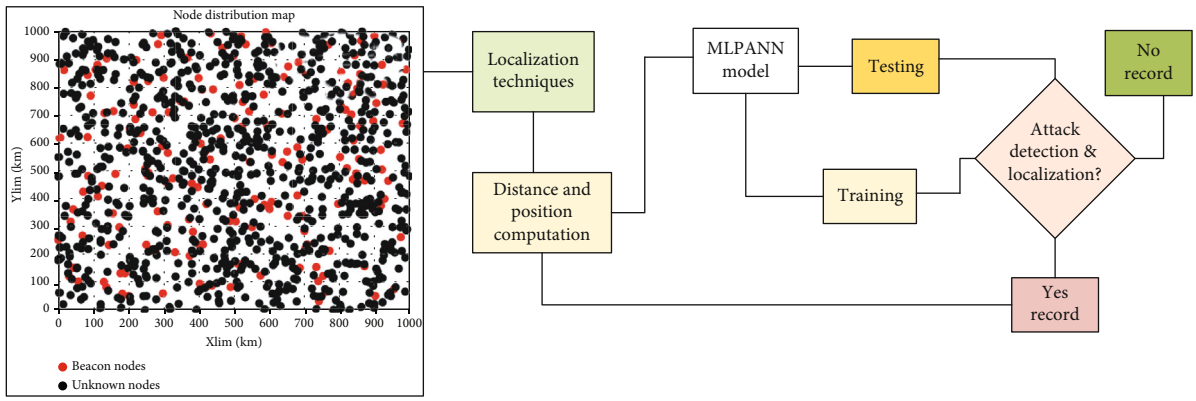


Figure 5: Secure localization techniques for detection and localization of malicious attacks using MLPANN in WSNs [36].

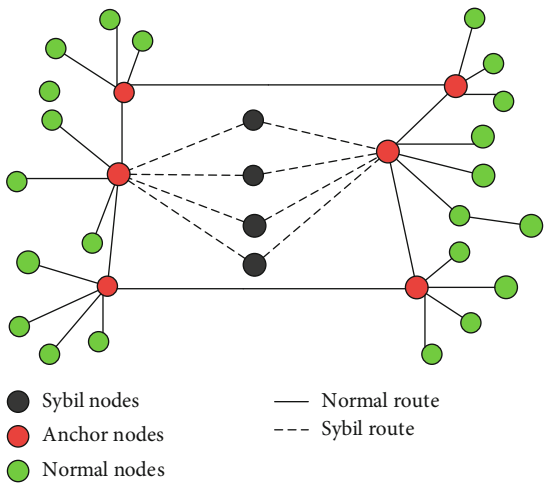


Figure 6: Illustration of routing Sybil attack in wireless sensor networks [2].

and localization. The data processing phase includes feature selection and normalization of the network traffic security dataset. The proposed system shown in Figure 7 is designed using optimized multilayer perceptron artificial neural network (MLPANN). The MLP is a feed-forward ANN with backpropagation to calculate the gradient used for weight calculation [48]. The ANN technique is a stochastic learning model for decision-making using interconnected information processing units [49]. ANN can estimate the nonlinear relationship between inputs and outputs and map the exchange of information among the nodes. The multilayer perceptron (MLP), as shown in Figure 8, configured with input layers, three hidden layers, and output layers. The proposed system used a gradient descent optimization for speeding and enhancing accuracy for detection and localization of attacks. This approach also uses a statically driven technique for training and testing using multilayer perceptron.

Several procedures are included in the proposed framework to identify malicious or unexpected routing. The method begins with a network data collection and preprocessing stage [50]. Next, it must find any missing values in the system and then fill in those blanks with appropriate values that were not present before processing began. We use the mean as our default. Subsequently, the dataset is cleaned up by removing any occurrences of duplicate values. After that, data encoding and normalization are carried out. In order to facilitate data handling, the encoded data undergoes a dimension's reduction procedure. To aid with anomaly detection, it is necessary to do feature optimization in order to extract the most useful characteristics from the data. In order to spot outliers in the dataset, optimal feature selection is crucial. For the same information, it aids in lowering the computational cost required to process it. Below is an equation that can be used to determine the entropy

$$E = -\sum_{i}^{L} P_i \log_2 P_i, \tag{13}$$

where p is the chance of finding a particular class label in the dataset. In this study, a hybrid machine learning approach is recommended for intrusion detection in a wireless sensor network after the optimal selection of features for anomaly detection.

4.1. Artificial Neural Networks. The multilayer perceptron artificial neural network (MLPANN) is a supervised machine learning approach using a human neuron model for data classification [47]. ANN processes and produces accurate information using a massive number of neurons and classify data-based neuron model that digests data and deliver correct output having layers and connecting nodes and active duty [51]. The layers of ANN are connected with nodes with the activation function. The configuration of typical ANN is depicted in Figure 9 with nodes and hidden layers varying from to three of the network using trainable parameters. ANN has wide applications in improving the efficiency of various schemes, including detection and localization of sensor nodes, routing and congestion control, and data aggregation in WSN. Artificial Neural Networks are data-driven tools for demonstrating nonlinear dynamic systems. They are efficient for the identification and modeling of nonlinear systems. They have standard approximation abilities and flexible structures to capture nonlinear

Table 4: Frequency distribution of DoS attacks for training and testing in the dataset [41].

Attack class	Frequency	Percent	Attack class	Frequency	Percent
Analysis	29	0.3	Generic	147	1.5
Normal	8632	86.0	Exploits	538	5.4
DoS	382	3.8	Reconnaissance	97	1.0
Backdoor	21	0.2	Shellcode	48	0.5
Fuzzers	122	1.2	Worms	17	0.2
Total	9186	91.5	Total	847	8.6

Table 5: Frequency distribution of DoS attack in WSN-DS and NSL-KDD dataset.

	WSN-DS dataset			NSL-KDD dataset	
Attack type	Frequency	Percent	Attack type	Frequency	Percent
Blackhole	2607	3.1	DoS	37403	33.7
Flooding	1019	1.2	Normal	61355	55.2
Grayhole	3287	3.9	Probes	10600	9.5
Normal	75300	89.1	R2L	913	.8
Scheduling	2343	2.8	U2R	839	.8
Total	84556	100.0	Total	111110	100.0

characteristics [52]. The input data $x1, x2, \cdots, xn$ denotes the input parameters of the dataset containing various protocols, services and the identity of the nodes, and $y1, y2, \cdots, yn$ represents the classified DoS attacks depending on the benchmark dataset. The ANN addresses the localization of sensor nodes and detection of the malicious nodes [18]. The proposed scheme has multiple trainable parameters for accurate attack localization and detection including the input nodes, hidden layers, bias and output nodes and also the connecting neurons. The ANN technique in WSNs improves the computational intelligence for scalable and adaptable features [30]. The ANN scheme was also used to obtain the accurate position of the sensor node using multilayer perceptron. It is also effective for prediction and clustering to get the location and position accuracy of nodes in WSNs.

The activation functions used for the artificial neural network multilayer perceptron are sigmoid and softmax, respectively. They are activation functions for the hidden layers and output layer, respectively. The sigmoid and softmax functions are stated below as in

$$Y = \frac{1}{1 + e^{-x}}, \tag{14}$$

$$z = \frac{e^{x}}{\sum_{k=1}^{n} e^{x}}, \tag{15}$$

where x is the vector of input to the output layer and y is the network's response with k index and n elements for the multilayer perceptron. The softmax function z is applied for activating for the classifier in the output layer. The number of hidden layers varies from one to three in our case for conducting the performance evaluation. The next step is to shrink the sampling dataset in order to better localize the feature that has to be extracted [53]. The pooling techniques are used to achieve this. In order to reduce the size of the image and the number of computations necessary, pooling is used. The max pooling approach was employed. Max pooling generates a new map after determining the feature maps' maximum value. A node's output in response to an input or combination of inputs is determined by the activation function of that node.

4.2. Optimization and Tuning Techniques. The goal of an ANN's optimization phase is to identify the optimal weighting scheme that leads to optimal performance. This is a tough optimization issue since it is categorized as a continuing nonlinear optimization problem. There are a lot of algorithms in books. Backpropagation is one of the most popular algorithms. This last one achieves excellent results, although it may run into a local minimum difficulty. To circumvent this issue and improve the likelihood of rapid convergence, we incorporate a local search method with a differential evolution algorithm. When training a neural network model, we utilize Adam optimizer to update the weights of the network based on the model's learned parameters with the greatest possible efficiency. Authors claim that Adam is a combination of the best features of two existing extensions of stochastic gradient descent: the Adaptive Gradient (AdaGrad) algorithm and the Root Mean Squared Propagation (RMSProp) algorithm. These two algorithms have a common characteristic: they both maintain a constant learning rate across all parameters. Adam sees the value in AdaGrad and RMSProp [53]. To fine-tune the weights of the neurons, we compute the gradient of the loss function and apply gradient descent optimization. The networks are trained via a gradient-based algorithm and the gradient descent nonlinear optimization method [29]. The gradient descent algorithm

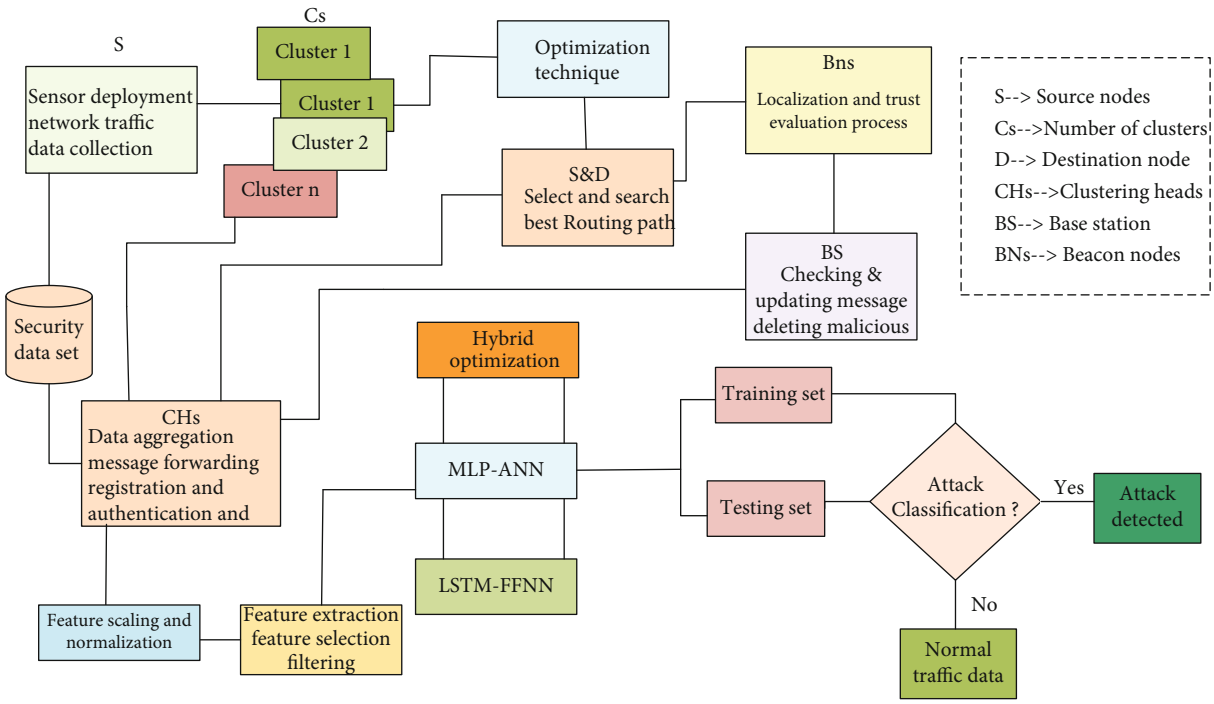


Figure 7: The proposed block diagram for attack localization and detection scheme in WSNs.

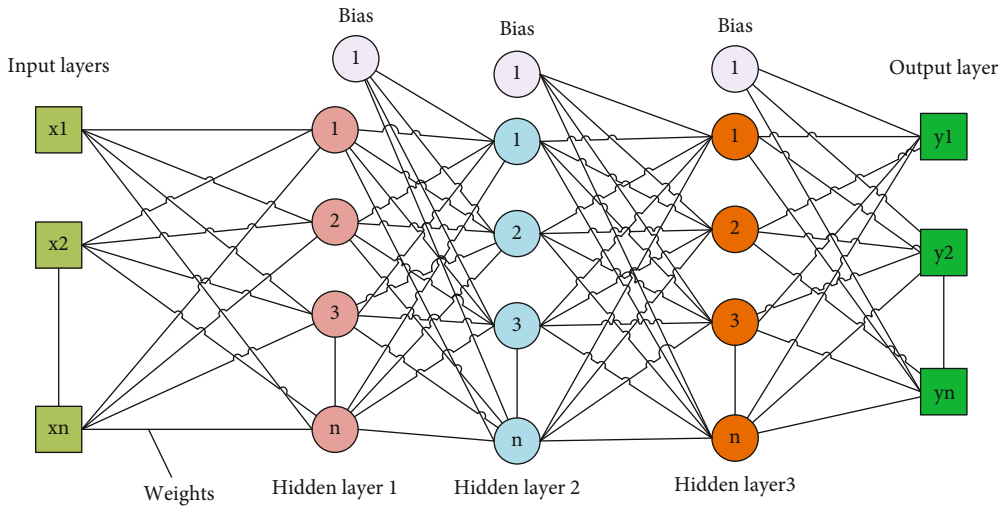
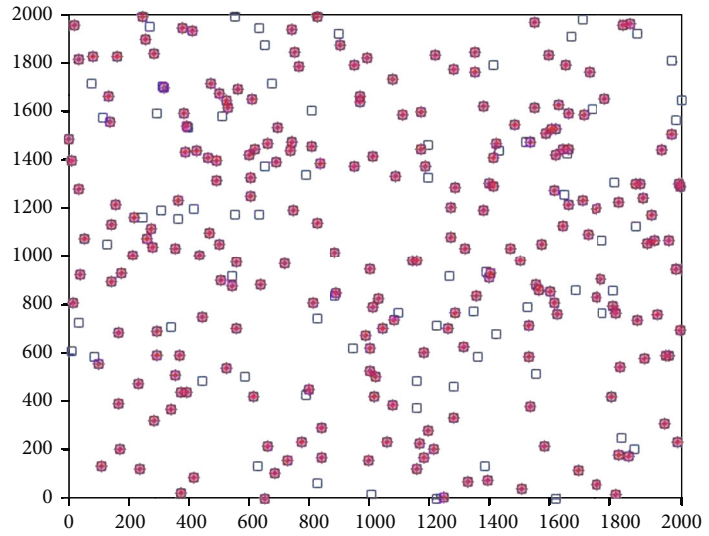


Figure 8: Structure of the proposed MLPANN architecture with three hidden layers [47].

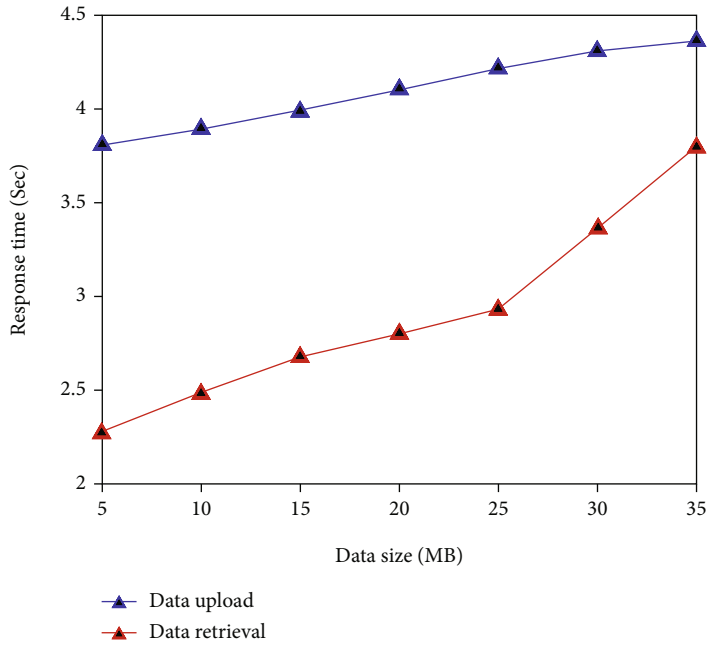
speeds up the training phase on the artificial neural network multilayer perceptron. The algorithm also helps to converge the weight iterations of the network.

Optimal weights, the optimal number of hidden layers and hidden nodes, and the optimal set of relevant characteristics are all necessary for building a multilayer perceptron [54]. The layer-by-layer weighted output data are collected at a secret node. The value of a bias node's weight is also given to it. One uses a nonlinear activation function on the aggregate of the weighted input values. The only restrictions are that the nonlinear function be differentiable and that the function's output values lie inside some interval. Finding an optimal set of weights that approximates both actual and estimated outputs is the goal of the MLP optimization issue. Continuous optimization is used to model this issue, invoking the problem classification of optimization techniques.

4.3. LSTM-FFNN. Using long short-term memory and feed-forward neural networks (LSTM-FFNNs), the suggested optimized multilayer perceptron artificial neural network achieved better results. Within the realm of DL, the long short-term memory (LSTM) RNN is analogous to a recursive function that repeatedly calls itself. With a recurrent neural network, the same computation is performed repeatedly on each data point in a recursive fashion, giving rise to the term "recurrent." The RNN suffers from the gradient vanishing and explosion issues. In contrast to other DL

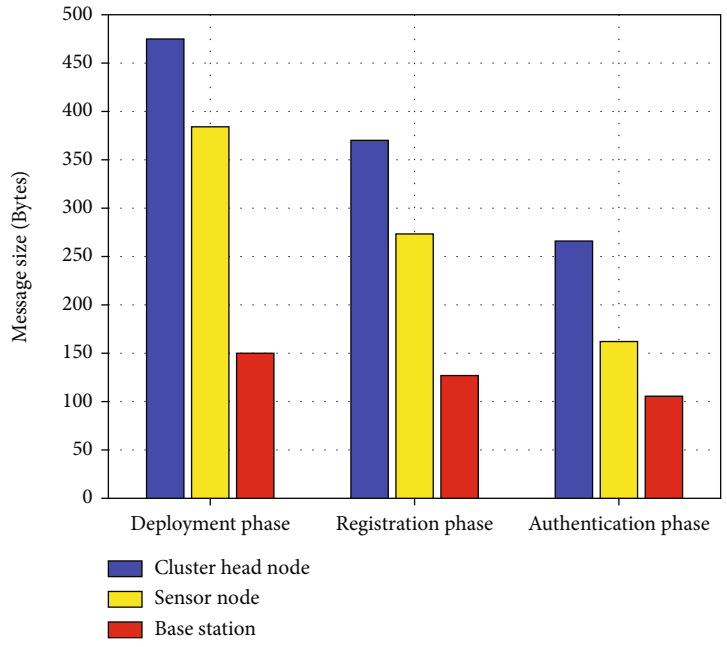


(a) Beacon node distribution phases

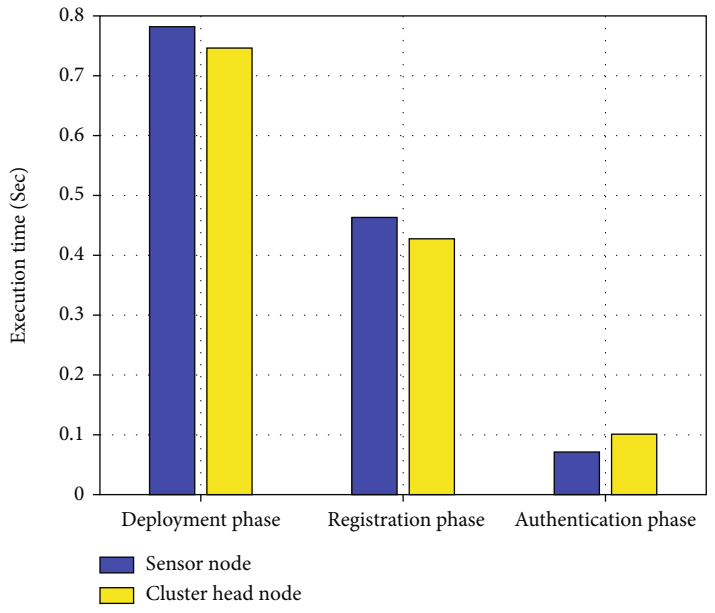


(b) Data uploading and retrieval phases

Figure 9: Continued.



(c) Authentication and registration phases in SN, CH, and BS



(d) Authentication and registration phases in SN and CH

Figure 9: The various phases of authentication and registration in the network model for secure data aggregation and transmission in WSNs.

techniques like deep NNs, the LSTM can recognize interdependencies in a time series and remember important data from earlier iterations to use in future predictions. We suppose that the model's inputs consist of the three preceding time steps. The data from the first unit flows into the second, as seen in the unfolded version [55, 56].

In contrast to the RNN-like LSTM, NNs like the fast forward neural network (FFNN) derive their predictions without looking back at prior time steps. They make their predictions based solely on data from the present lag. Inputs plus an n-node hidden layer make up the FFNN. Each node's output is a function of its inputs and the weights of the connections between them. Our model consists of five distinct layers: a vector input layer, three hidden layers, and a single-node output layer that returns a 1 or 0 depending on the type of classification being performed.

In this paper, we applied various activation functions considering the threshold value. The ReLU activation function is being used. It is an acronym for a nonlinear operation's rectified linear unit. Since the real world is typically highly nonlinear, the goal of adding nonlinearity to the network is achieved. It can be defined as mathematically as in

$$\text{Re Lu}(x) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{16}$$

The sigmoid (or logistic) activation function (δ) was used to input the z value into the logistic function and generate values between 0 and 1 using a threshold of 0.5 as the reference value. This can be defined as mathematically

$$\delta(z) = \frac{1}{1 + e^{-x}}. \tag{17}$$

The choice of activation function is motivated by the presence of two classes of labels (outputs) for this method. Therefore, the method of binary classification technique should be used. Cross-entropy, a well-known loss function for ANNs, was the one we employed. Specifically, [47] defines cross-entropy (C) as

$$\begin{aligned} C &= -\frac{1}{n}\sum_{x}^{n}[\text{ylna} + (1-y)\ \ln\ (1-a)], \\ a &= \delta(z) = \delta\left(\sum_{j} w_j x_j + b\right), \end{aligned} \tag{18}$$

where Adam, the hybrid optimizer, will iteratively adjust the weights w and biases b. Adam is an improved version of the Stochastic Gradient Descent (SGD) algorithm. According to the scikit-learn documentation, Adam performs reasonably well on huge datasets. There are four variables you can adjust in Adam: the rate at which one is learning, the exponential decay rate for first-moment estimates, the rate at which one's second-moment estimates decay, and a very small amount to avoid a division by zero. Also, Adam is superior to SGD in noisy environments because it combines the advantages of two other popular optimizers (the adaptive gradient algorithm and root mean square propagation). We get things off with an early architecture for hyperparameter optimization that seeks optimal performance with as little computational complexity as possible.

5. Simulation and Result Discussion

The simulation setting configuration and evaluation metrics will be discussed in this section. Wireless sensors are distributed randomly forming clustering with cluster heads in the target field with an area of $1000 \times 1000\ \text{m}^2$. The routing protocols are used for making clustering and selection of the cluster head in each round of the simulation and localization of the unknown nodes with help of the beacon nodes and sink nodes. The cluster head achieves more computational data processing from the sensor nodes and communication with base station. The simulation parameter configuration is shown in Table 6. Intel (R) Xeon (R) Silver 4214 CPU @ 2.20GHz 2.19GHz (2 processors) with 128 GB (128 GB useable), x64-based processor, and 64-bit operating system running Windows using MATLAB R2021a is used for network planning and simulation.

Table 6: Simulation setup for the proposed network model.

Parameter	Values
Number of sensors	300-1000
Beacon nodes	60-120
Unknown nodes	240-840
Protocol type	Clustering and routing
Deployment area	$1000 \times 1000\ \text{m}^2$
Mobility	Random
Number of clusters	10
Sink position	500, 1000
Number of attacks	5-60
Data size	4000 kb
Attacks	Routing
Transmission radius	400 m

Our primary effort is devoted to determining how well various hybrid-based improvements to the original DV-hop algorithm perform in detecting and pinpointing hostile nodes that have hijacked the beacon node and are supplying false routing information [57]. All of our proposed algorithms have been implemented in the MATLAB simulator for thorough testing and analysis of their localization faults and precision. Numerous researchers rely on MATLAB, a simulation programed and numerical computing environment, to test out new ideas, conduct research, and build models. In our tests, we have examined the localization accuracy and the localization error per node by changing the percentage of anchor nodes, the total number of sensor nodes, and the nodes' communication range across four different topologies. One way to measure an algorithm's efficacy in localization is by looking at how it performs on average with regard to localization errors. We employ IBM SPSS, Python, and the WEKA Java toolboxes for data processing and analysis to gauge the effectiveness of the suggested strategy against the dataset [58]. The average error of localization to all the nodes is calculated using Equation (15). The clustering and routing protocols are used for clustering and selection of the cluster head selection and maximizing the network lifetime and improving the network performance. The routing attacks including the sinkhole attacks, blackhole attacks, and Sybil attacks are used in the simulation scenario for evaluating the localization and detection accuracy.

The simulation results depict that the data processed from the environment is authenticated and registered. Figure 9 shows data processing and aggregation by the cluster head sent to the base station (BS). Figures 9(a) and 9 show the dynamic clustering and data retrieval of the sensors by the beacon nodes. The cluster head (CH) aggregates huge message size as in Figures 9(c) and 2(d); the sensor nodes (SNs) consume greater time form data execution.

Registration phases are utilized to identify sensor nodes, aggregation nodes, and base stations using smart contact of the public blockchain [59]. The intelligent communications verify the existence of the aggregation node validated by its MAC address and its identity checked by the base station.

The public blockchain records of validated aggregated nodes and stored data of the aggregated node provide reliable authentication techniques in WSNs. The sensor nodes are allowed to join the blockchain after the completion of the registration process to reduce external attacks on WSNs. The sensor nodes have aggregation nodes after random deployment in the target field. The aggregation nodes authenticate the identities of the sensor nodes using a private for communicating with them, and the base station also authenticates the aggregation node for communicating with it using a public key. The aggregation nodes communicate with each other using mutual authentication process.

Figure 10 shows the distribution and the experimental simulation of the nodes. Moreover, this work introduces the average localization error and coverage, localization, and detection accuracy as evaluation metrics. The average localization error (ALE), average localization accuracy (ALA), accuracy, detection rate precision, and recall are used as evaluation metrics. The average error localization, shortened as ALE [2], is computed as follows in Equation (19). The ALE is the summation of the LE of all the unknown nodes to the total number of unknown nodes. The LE is the difference between estimated and actual position of unknown nodes.

$$\begin{aligned}
\text{Localization Error (LE)} &= \sqrt{\left(u_i' - u_i\right)^2 + (v_i' - v_i)^2}, \\
\text{Average Localization Error (ALE)} &= \sum_{i=1}^{n} \frac{\sqrt{\left(u_i' - u_i\right)^2 + (v_i' - v_i)^2}}{\text{nR}}, \\
\text{Average Localization Accuracy (ALA)} &= \left(1 - \left(\sum_{i=1}^{n} \frac{\sqrt{\left(u_i' - u_i\right)^2 + (v_i' - v_i)^2}}{\text{nR}}\right)\right) \times 100\%,
\end{aligned} \tag{19}$$

where (u_i' and v_i') are the real coordinates of the anonymous node i and (ui, vi) are the computed coordinates, n denotes unknown nodes, and R is radius of communication in the network. Wireless sensor nodes are deployed and simulated using localization process using the beacon nodes as in Figure 10(a). The error for the anonymous sensor nodes is displayed in Figure 10(b). The position and error for each node are computed using the localization scheme. The computation of the localization accuracy for each node enables for effective identification and localization of the malicious nodes with help of the beacon nodes and the base station.

The effectiveness of the distance vector hop algorithm is measured by malicious node detection, localization accuracy, and localization efficiency [60]. The practical localization estimation of the unknown and malicious nodes is determined by the number of the anchor nodes for its evaluation metrics, as shown in Figure 11. The relative error defines between the computed position of the node and the actual location of the node. Malicious nodes affect nodes' distribution and localization accuracy by creating the wrong position and location of the unknown sensor nodes in WSNs. Malicious nodes mislead the sensor nodes' routing path and information, making the network service and performance degrade.

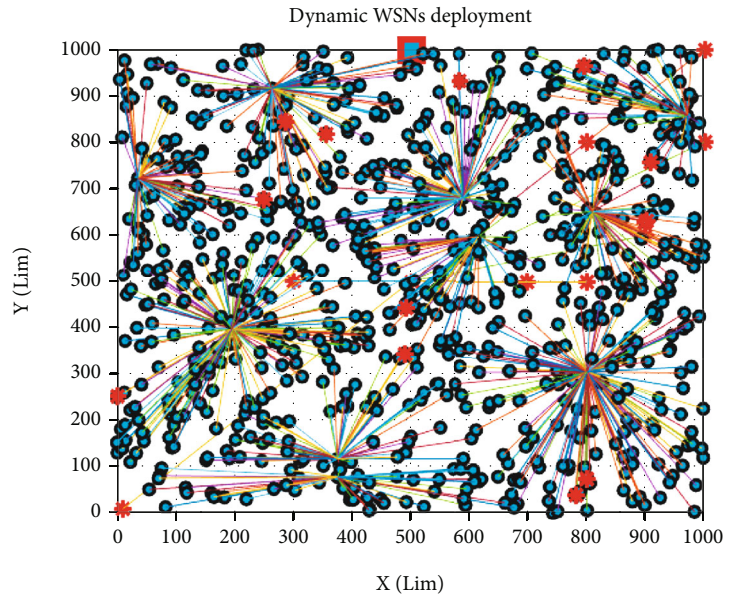
The average localization accuracy and detection accuracy of the proposed system are 99.51% and 99.83% with 840 unknown nodes and 160 beacon nodes for accurate computation of malicious nodes, respectively, as shown in Figures 11(a) and 11(b).

According to the findings of the simulation, anchor nodes have a greater number of neighbors and a higher degree of connectedness than regular sensor nodes, as can be seen in Figure 12(a). If we use the regular model, we can determine that the average connectivity of the network is 404, and the average number of neighbor nodes that each anchor node has is 63. As can be seen in Figure 12(b), the overall network's average localization error was reduced to 0.0049 thanks to the simulation's efforts, and this was achieved across all nodes. This would imply that all sensor nodes are precisely located and have a unique identity thanks to the beacon nodes, which help in the identification and localization of malicious nodes.

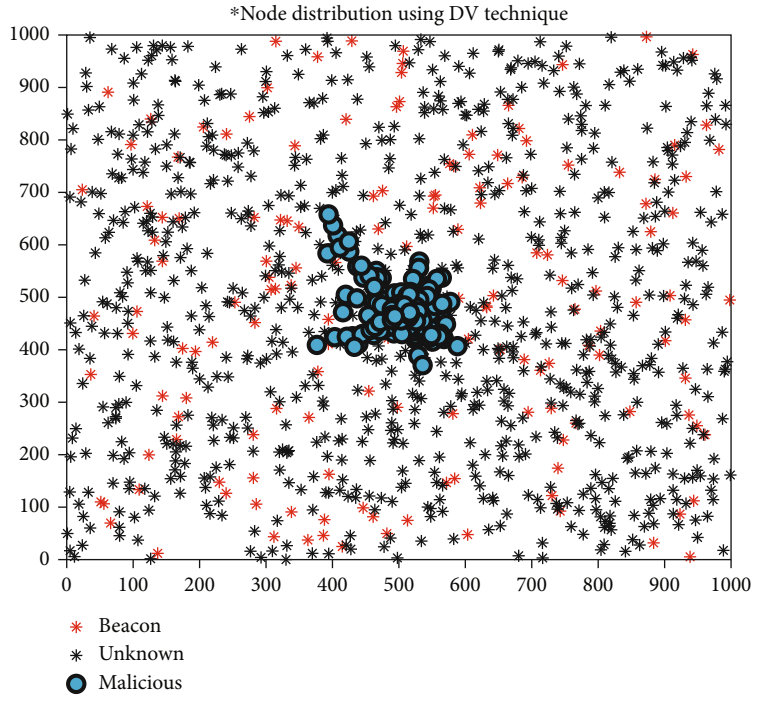
The simulation results demonstrate that the suggested method utilizes hybrid localization techniques utilizing both range-free and range-based approaches to accurately determine the position and location of each unknown node while minimizing energy consumption. As twenty mobile anchor nodes are utilized in the proposed method, the price is kept low while the accuracy of pinpointing malicious nodes in WSNs is much enhanced. Figures 13(a) and 13(b) for the beacon nodes and the unknown nodes, respectively, illustrate how the hybrid strategy combining the DV-hop technique with other approaches such as RSSI and DE improves the average localization accuracy of the proposed scheme.

The experimental findings for calculating the location error against changing numbers of beacon nodes are displayed in Figures 14(a) and 14(b). In addition, the localization error for all the algorithms gradually decreases as the number of the activated sensor nodes grows [34]. The proposed hybrid method has the lowest localization error score of all the methods we have tested. With 200 beacon nodes, more reference points are detected, reducing the margin of error for localization. Figure 14 shows conclusive proof that the new method outperforms conventional location-based algorithms when it comes to pinpointing the origin of an error. In the same setup, nearly all of the methods that have been tried and tested have been effective. As a result of having more points of reference for the target nodes, the suggested method allows for a gradual decline. In contrast, the network is strengthened by an adequate number of anchor nodes, as the distance between the unknown nodes and the anchor nodes decreases.

As may be shown in Figures 14(a) and 14(b), the ALE of four different localization techniques decreases as the number of beacon nodes increases. Since there are more anchor nodes now, the average distance travelled in one hop can be calculated with more precision. The distances predicted by the anchor nodes from the unknown nodes are more accurate [61, 62]. This shows that the proposed approach is effective to estimate the placement of unknown nodes as the number of anchors grows because it has more



(a) Clustering and localization of WSNs

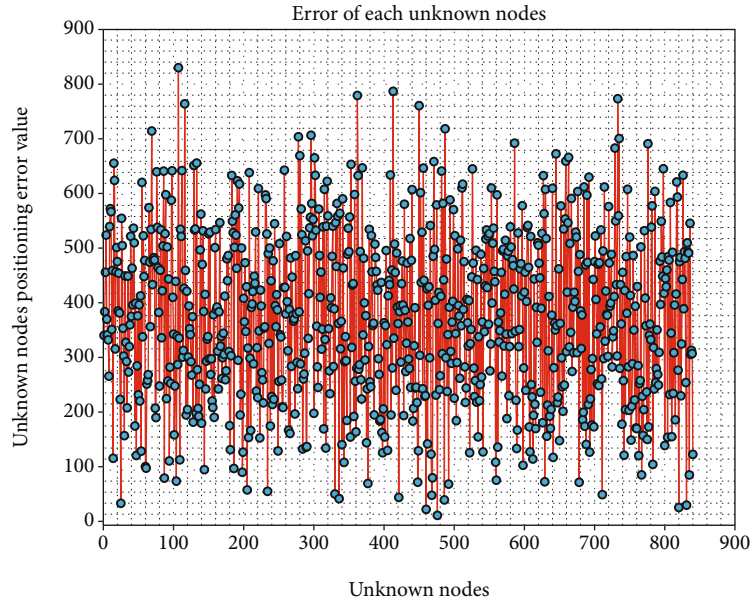


(b) Malicious node localization in WSNs

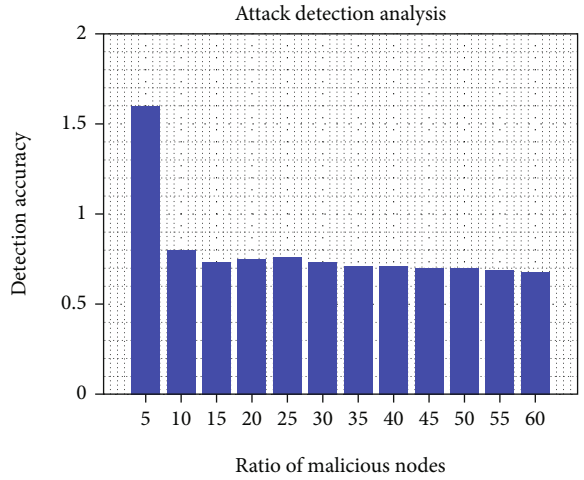
Figure 10: Sensor node deployment using distance vector protocol and triangulation process in WSNs.

circumstances to work with. Given that some fraction of the nodes can serve as anchor nodes for node localization, the suggested methodology exhibits lower error compared to the previous methods.

5.1. Performance Metrics. In this section, a cybersecurity dataset is applied with different types of attack categories. The benchmark datasets are utilized for analyzing and processing using the optimized artificial neural network technique to detect and localize multiple attacks to evaluate the proposed system about the Sybil attacks. The dataset is used as a benchmark for the security localization and detection of accuracy of various classes of routing attacks in the network. The Python programming language, SPSS, and WEKA toolboxes are used for data processing and classification to detect the different classes of attacks in WSNs [20]. There are different types of performance metrics for measuring the effectiveness of the proposed scheme. Table 7 shows the measurement entities of the system with mathematical equations.



(a) Positioning error of unknown nodes

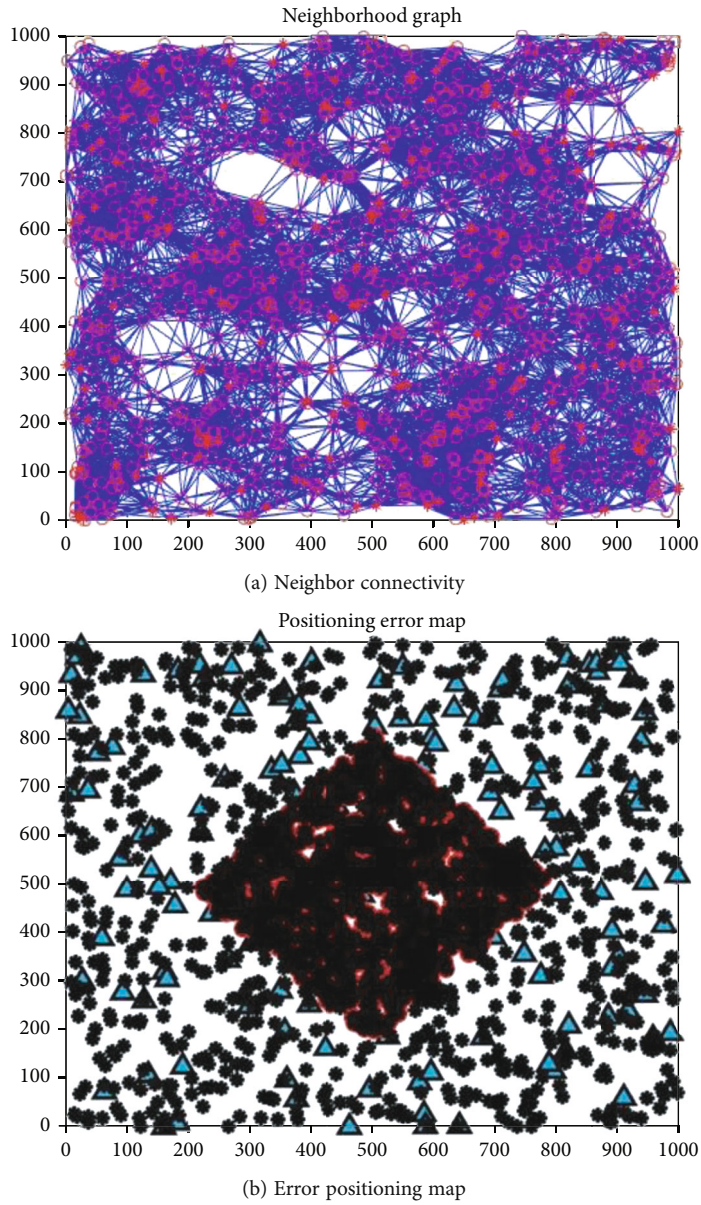


(b) Detection of accuracy of malicious nodes

Figure 11: Computing localization and position of unknown node errors and detection accuracy of malicious nodes using distance vector hop algorithm and trilateration process.

The parameters are false negatives = F_n [47], true positives = T_p, true negatives = T_n, and false positives = F_p [63]. Accuracy is the parameter for evaluating the performance of proposed classification model [64]. Informally, it is the section of predictions the model achieved successfully. Formally, it can be computed [20] as shown in Table 7. The F-measure [65] is a combination of recall and precision [63] computed as in Table 7. The Matthews correlation coefficient (MCC) is also the measure of the performance for scoring prediction of the model. The proposed system has training and testing phases using the cybersecurity dataset as benchmark with different classes of attacks. The system is based on the ANN approach achieves an accuracy of 99.84% and error of 0.16% using three hidden layers with 10-fold cross-validation using CICIDS2018 a benchmark dataset. The different attacks are correctly detected and localized, greater than 78% proposed by Dong et al. [2] using the distance vector hop scheme with an error of 22% malicious node localization.

Receiver operating characteristic analysis is useful to assess the model's accuracy using the ANN technique [66]. The total area under the ROC represents the statistical probability prediction of the classification of the proposed model for different types of attacks using a threshold cutting point $C\epsilon(0, 1)$ as shown in Table 7. This ROC analysis supports the inference area under the curve and Precision-recall curves. The ROC is a plot of the sensitivity versus 1-specificity, as shown above in Figure 15(a). Sensitivity is the number of attacks correctly identified in the network. 1-specificity is the attack classes wrongly rejected. The cumulative chart gain in Figure 15(b) shows the overall percentage of the total observations for the given class of attacks in the network. The target category is the percentage of the overall amount



(a) Neighbor connectivity

(b) Error positioning map

Figure 12: Simulated results of sensor deployment (a) and neighbor relationship diagram (b).

of samples in the dataset. The diagonal line is the baseline for the classification of the target samples. The cumulative gain chart is a cutoff choosing the attack classification and mapping the appropriate cutoff values. Table 8 depicts the area under the curve detection rate performance for each type of attack.

The area under curve is visualized in Figure 16 for each attack category of the dataset. The area under the curve is more significant for the standard class in the proposed network traffic analysis. The ROC analysis shows that the MLPANN approach is practical for multiclass attack localization and detecting DoS attacks. The ROC shows that the proposed scheme is effective for DoS attack classification using a benchmark dataset.

The area under the curve is a statistical summary of the ROC curve, and the values represent each attack category. The area under the curve also indicates the probability of the classification model. The standard class of the attack has a greater extent, which is effectively detected. The pseudopredictive probability in Figure 17 describes each attack class's scaling by dividing their sum of classification accuracy. The effectiveness of the scheme can be proved using other attacks types and datasets as threshold measurements.

The average classification of the proposed system is 96% using the predictive probability model. Figure 17 shows the effectiveness of the various attack classification model. The classification model is trained 80% of the dataset, and 20% tested samples using the batch mode with the activation function using gradient descent algorithm with three hidden layers and trainable parameters.

5.2. Performance Comparison. The performance of the proposed methodology is validated and confirmed by comparing and testing with other previous works using various

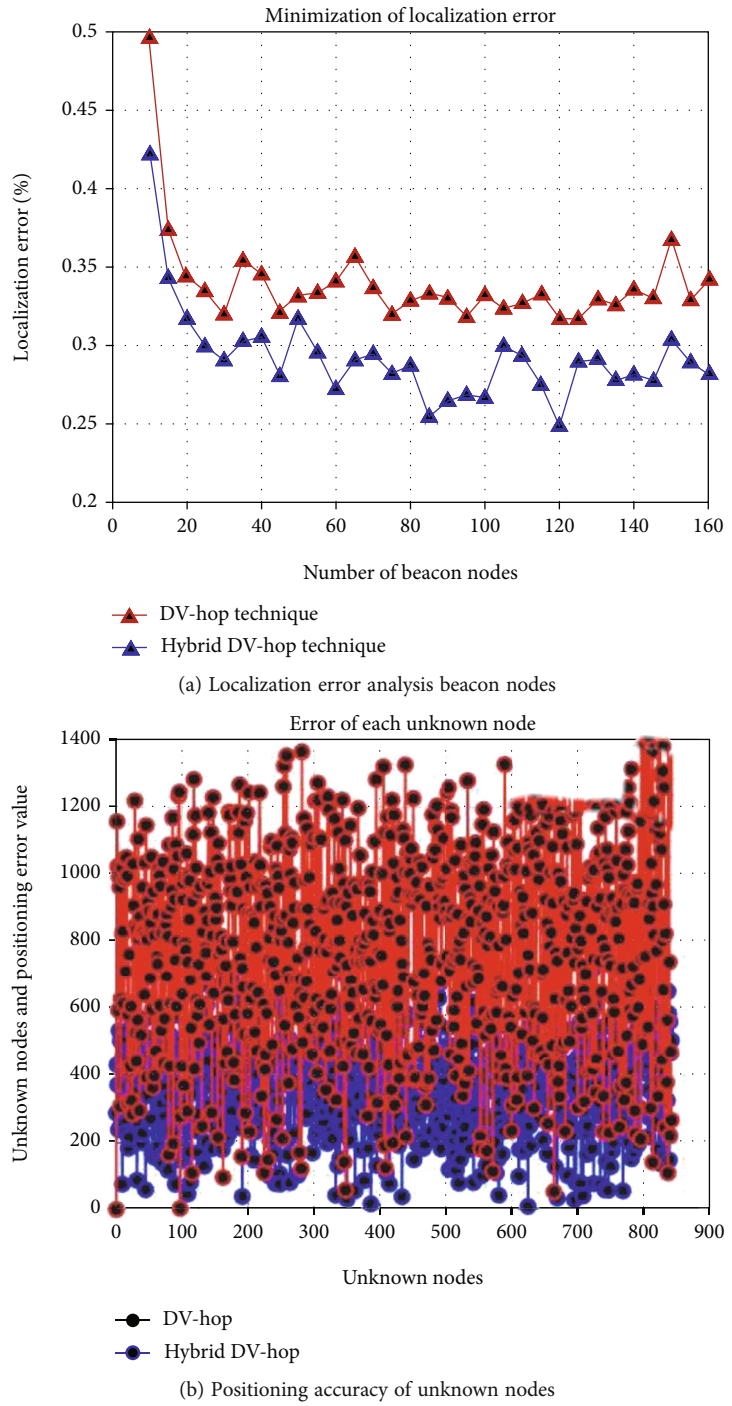
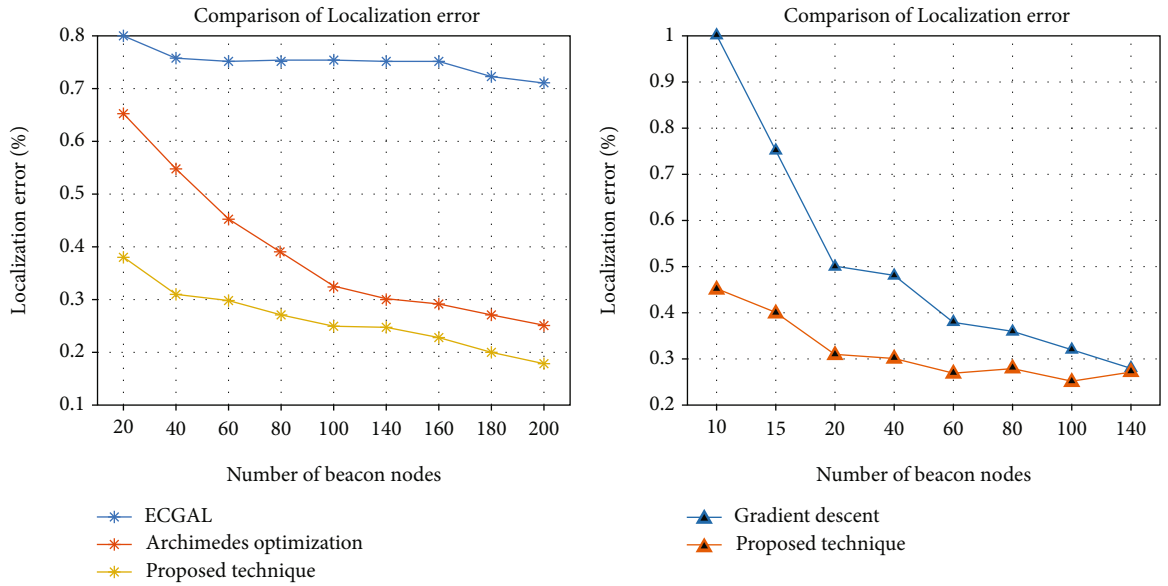


Figure 13: Improving the localization and position accuracy of wireless sensor nodes using hybrid scheme.

benchmark datasets. The comparison performance of recent works as shown in Table 9 suggests the optimized MLPANN technique is effective for detection and localization of attacks in WSNs.

Figures 18(a) and 18(b) show the performance comparison of the proposed system using four benchmark datasets and different attack detection models using the accuracy, precision, recall, and F1-measure. This suggests that the proposed scheme is effective for detecting and localization attacks in WSNs.

This suggests that the proposed system is more effective than the previous work by Almomani et al. [20] artificial neural network-based intrusion detection system (ANN-IDS) for routing attack detection and classification, as shown in Figure 18(b) with an average detection accuracy of 97.2% using sample WSN-dataset using ten-fold cross-validation with three hidden layers. Dong et al. [2] used the distance vector hop algorithm to detect Sybil attacks with a localization accuracy of 78%, which is less than the proposed scheme. The proposed work is also



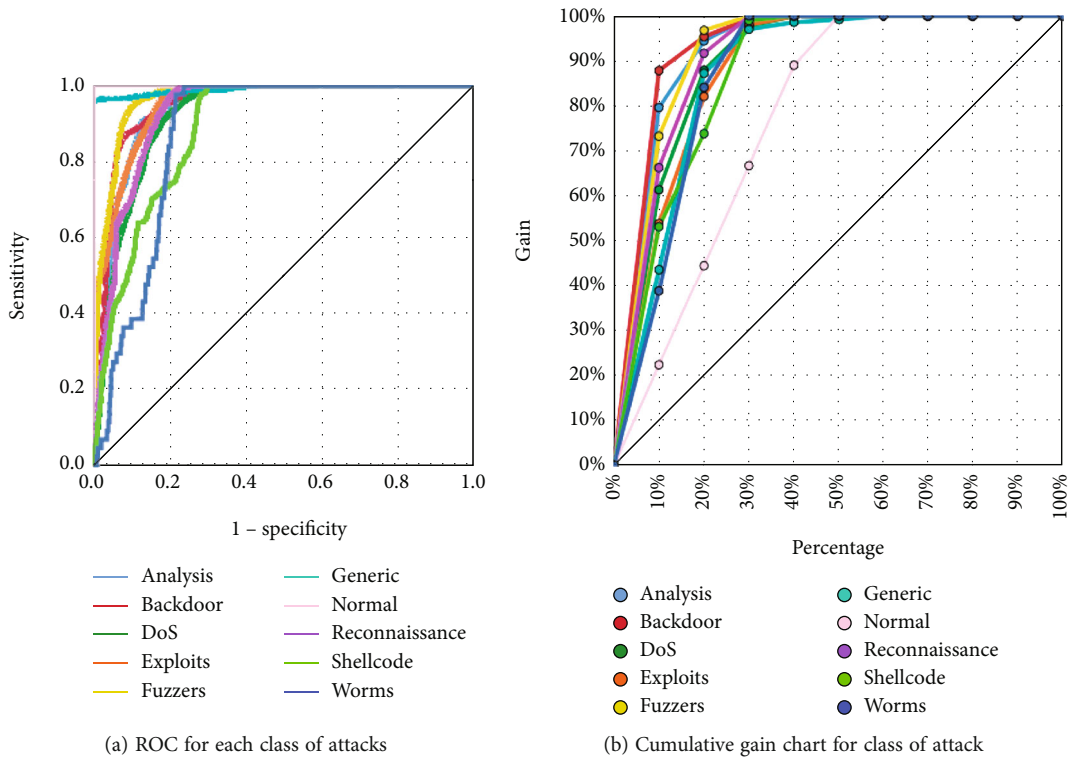
(a) Comparison of localization schemes

(b) Comparison with gradient descent

Figure 14: Comparison of the localization error of the proposed scheme with other modes by varying the number of nodes.

Table 7: Performances metrics of the proposed system with their technical and mathematical expressions.

Metrics	Technical description	Mathematical equations
Sensitivity	The positive prediction of the model	$\text{Sensitivity} = \frac{T_p}{T_p + F_n}$
Specificity	The negative prediction of the model	$\text{Specificity} = \frac{T_n}{F_n + T_p}$
Precision	Truly identified instances of samples	$\text{Precision}(p) = \frac{T_p}{F_p + T_p}$
ROC	Classifier performances of the model	$\text{ROC}(.) = \{\text{FPR}(C), \text{TPR}(C), C\epsilon(0,1)\}$
Recall	Equivalent to TP rate	$\text{Recall} = \frac{T_p}{T_p + F_n}$
F-measure	Combination of precision and recall	$\text{F} - \text{measure} = \frac{2 \times T_p}{2 \times T_p + F_p + F_n}$
Accuracy	Classification prediction of the model	$\text{Acc.} = \frac{T_n + T_p}{T_n + T_p + F_n + F_p}$
FP rate	Wrongly classified events	$\text{FPR} = \frac{F_p}{F_p + T_n}$
Energy	Energy consumption analysis	$E_c = E_t - E_l$
Lifetime	Over all alive nodes (N)	$\text{Aliv}(N) = \text{Total}(N) - \text{Dead}(N)$
MCC	Binary classification	$\text{Mcc} = \frac{T_p.T_n - F_p.F_n}{\sqrt{(T_p + F_p).(T_p + F_n).(T_n + F_p).(T_n + F_n)}}$
ALE	Average localization error	$\text{ALE} = \sum_{i=1}^{n} \frac{\sqrt{\left(u_i' - u_i\right)^2 + (v_i' - v_i)^2}}{nR}$



(a) ROC for each class of attacks

(b) Cumulative gain chart for class of attack

Figure 15: Receiver operating characteristics (ROCs) and cumulative chart of the proposed system.

Table 8: Area under the curve for each class of attack.

Class	Analysis	DoS	Backdoor	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Shellcode	Worms
Area	99.8%	98.8%	97.2%	99.1%	99.9%	96.7%	100%	99.2%	97.8%	94.9%

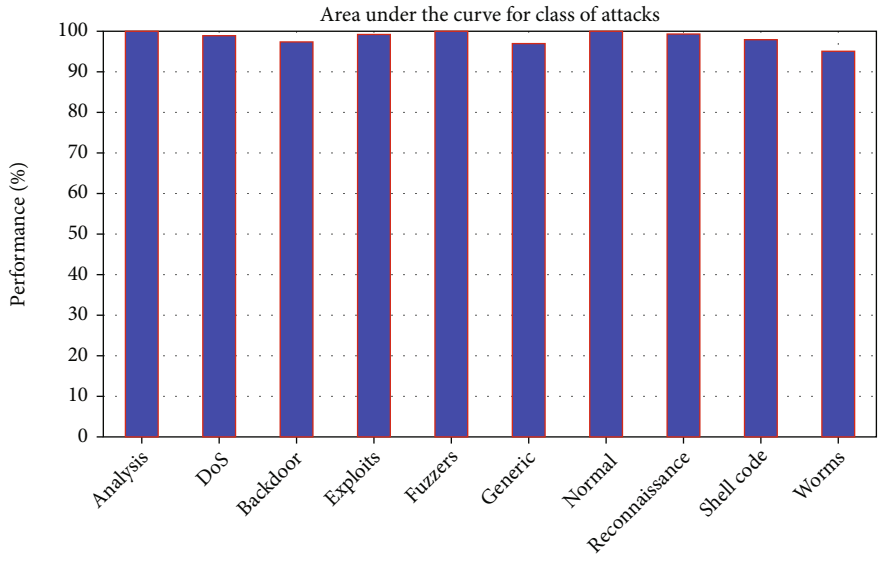


Figure 16: Area of under the curve for each class of attack using UNSW-NB 15 benchmark dataset.

practical compared to the MK-ELM model [67], which has an accuracy of 92.10% using UNSW-NB 15 dataset. Figure 18 shows the detection and localization for the proposed ANN approach compared with other works for Sybil attack detection. The comparison performance is using sample experimental dataset examined by Sujatha and Anita [24] with an average detection rate of 97% using fuzzy extreme machines (FEMs).

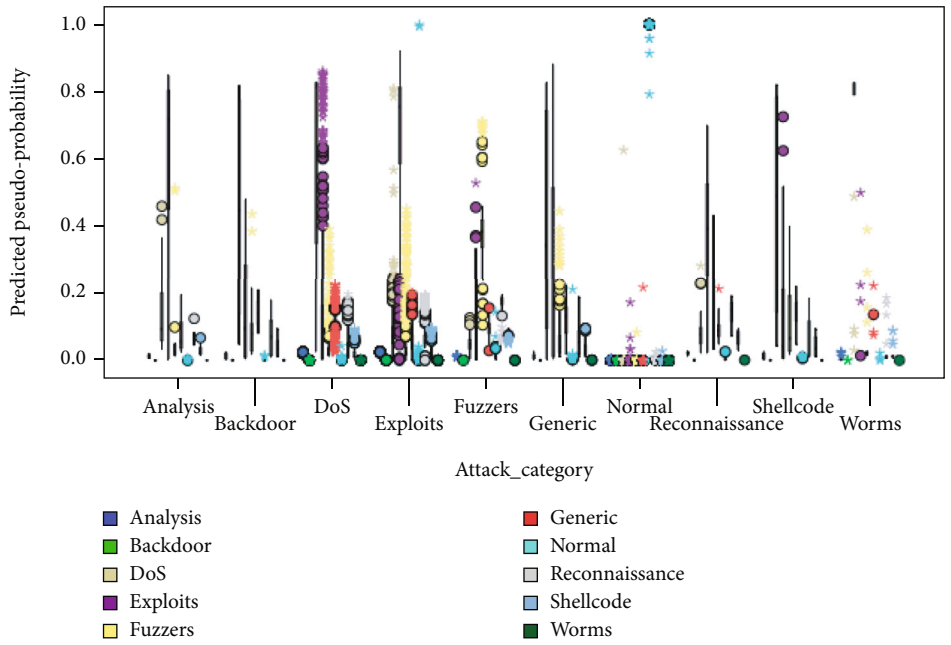


Figure 17: Attack classification prediction probability of the model.

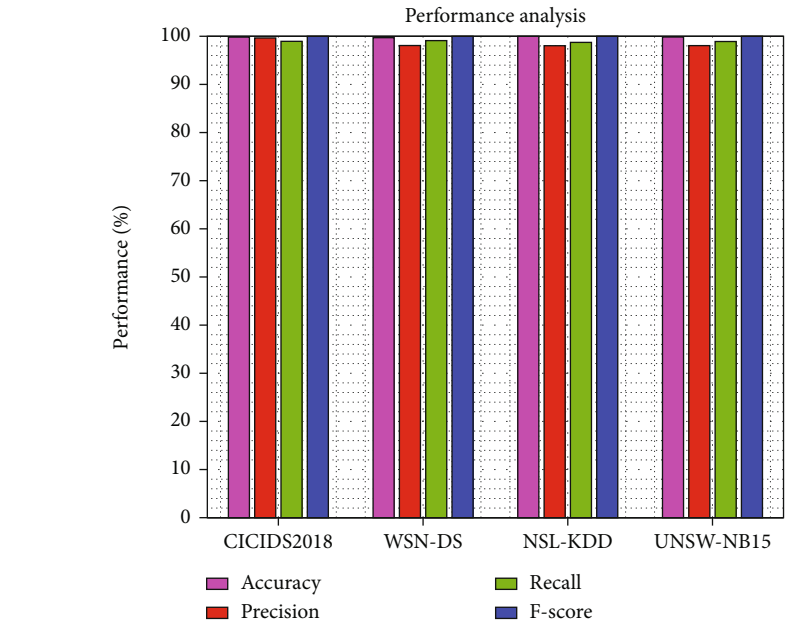
Table 9: Comparative analysis of the proposed technique with recent attack detection models using CICIDS2018 dataset.

	Methods	Accuracy	Precision	Recall	F1-score
Zhou et al. [8]	MLP	97.56	99.12	98.79	98.23
Almomani et al. [20]	ANN-IDS	97.18	97.8	97.5	97.69
Sujatha and Anita [24]	FEM	99	98	98	98.98
Zhang et al. [67]	MK-ELM	98.34	98.03	97.63	97.64
Khan et al. [68]	LEACH++-ANN	97	96.8	96	96.4
Mohammadpour et al. [69]	CNN-MCL	99.46	99.76	99.15	99.46
Xinlong and Zhibin [70]	HTM-LSTM	97.74	97.20	97.92	97.72
Proposed system	MLPANN	99.83	99.71	100.00	99.85

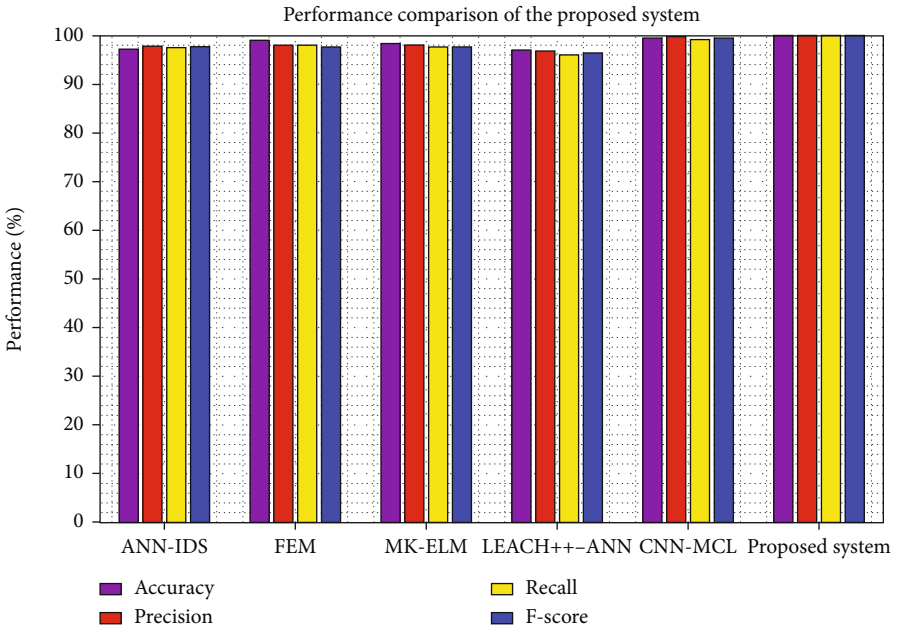
The proposed attack detection and localization scheme achieve 100% using the same dataset. Hasan et al. [71] determined the detection accuracy of 91.66% of the malicious node using an optimized artificial neural network using the packet delivery and energy consumption evaluation metrics. The various comparison performances conclude our proposed scheme is effective for the detection localization of attacks in WSNs. Khan et al. [68] analyzed the detection of routing attacks using the LEACH++ protocol based on an artificial neural network (LEACH++-ANN) and achieved a detection accuracy of 98%. This proves the proposed scheme is more effective for detecting routing attacks, with an average detection accuracy of 99.62%. The proposed system also achieves average detection accuracy of 98.4% using the benchmark dataset NSL-KDD as shown in Figure 18(b) for each class of attack. The proposed approach is practical for detecting and localization DoS attacks in WSNs compared to the convolutional neural network and mean convolutional layer (CNN-MCL) model proposed by Mohammadpour et al. [69] with an average detection accuracy of 99.46%. Zhang et al. [67] proposed an hierarchical intrusion detection model (HIDM) for WSNs using a multikernel-based extreme learning machine (MK-ELM) classification technique using UNSW-NB and NSL-KDD benchmark datasets.

The proposed system's average localization and detection rate are validated by comparing previous works with different classes of attacks. Table 10 shows that when applied the UNSW-NB 15 dataset, which serves as a benchmark for identifying and classifying routing assaults, the proposed approach improves detection accuracy by class using 80% of training and 20% of testing of samples with five hidden layers. The demonstration further shows that the verification of the suggested performance parameters and metrics (accuracy, precision, F1-score, and recall) against those of recently published attack detection models.

The performance of the proposed system is effective for detection and localization of DoS attacks in WSNs using benchmark datasets in terms of the evaluation metrics such as accuracy, precision, recall and F1-score as shown in



(a) Comparison using benchmark datasets



(b) Comparison of various attack detection models

Figure 18: Performance comparison of the proposed system using benchmark datasets for attack detection and classification.

Table 10: Comparative analysis of the proposed technique with recent attack detection models using UNSW-NB 15 dataset.

Author	Method	Accuracy	Precision	Recall	F1-score
Pasikhani,et al. [72]	RL-IDS	98.35	98.36	97.04	98.34
Upadhyay et al. [73]	GBFS-IDS	92.96	92.50	92.40	92.44
Abdan and H. Seno [11]	ML-ID	98.9	87.7	99.6	92.78
Gudla et al. [74]	DI-ADS	99.44	99.02	99.60	99.30
Alghamdi [75]	PO-CFNN	99.86	99.89	99.58	99.72
Khilar et al. [76]	DNN-CSO	99.46	99.75	99.62	99.76
Proposed system	MLPANN	100.00	100.00	100.00	100.00

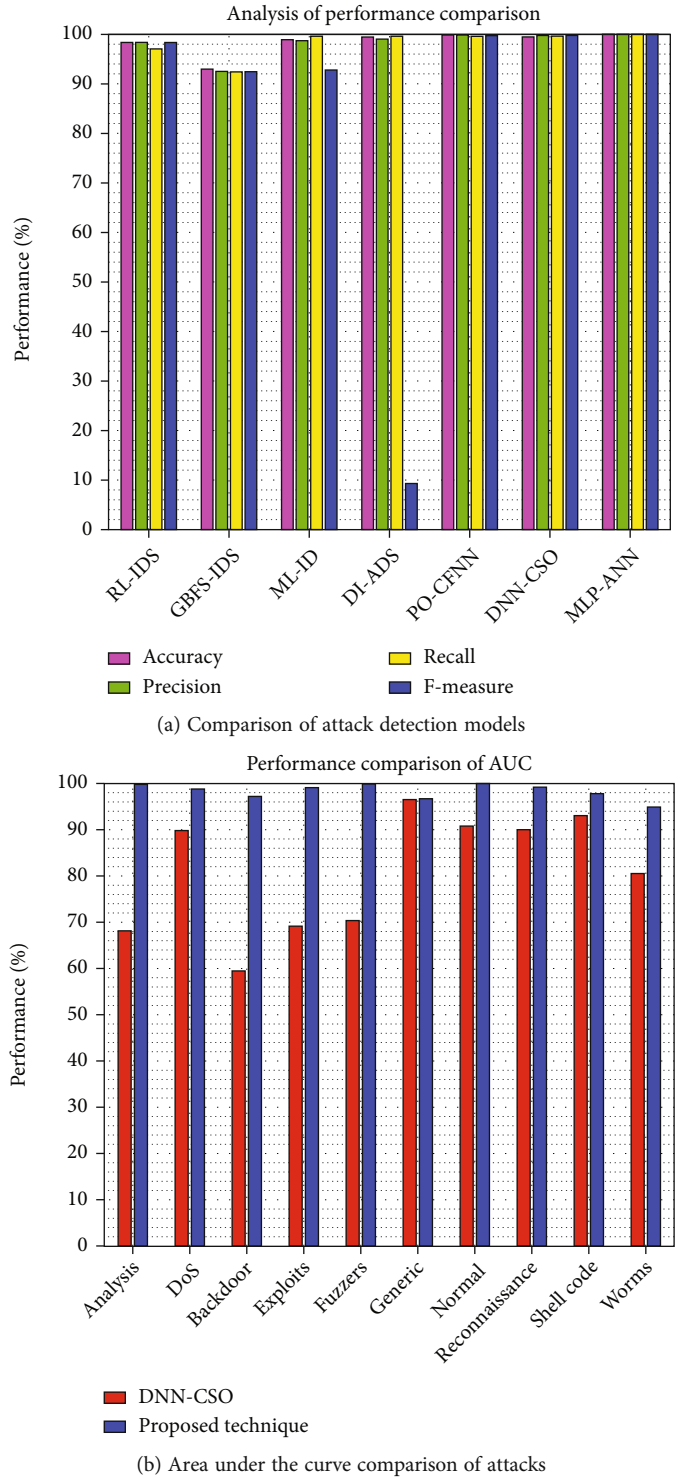


Figure 19: Performance comparison of the proposed system with recent works using benchmark datasets.

Figure 19(a). The area under the curve (AUC) is also used for evaluating the performance of the system as show in Figure 19(b). This confirms that the optimized MLPANN approach is effective in attack detection and localization of WSNs attacks.

The proposed multilayer perception artificial neural network (MLPANN) technique is further compared with MK-ELM using the NSL-KDD benchmark dataset, taking a section of 14,000 sample records with three hidden layers, as shown below in Table 11. The average detection accuracy of the proposed technique is 98.4% using 111,110 samples which is more effective than MK-ELM with 14,000 samples with an average detection accuracy of 98.34%.

Table 11: Comparison of performance of the proposed MLPANN with MK-ELM using the section of the NSL-KDD dataset with hidden layers.

Class of attacks	TP rate (%)		FP rate (%)		FN rate (%)		TN rate (%)	
	MK-ELM	Proposed scheme	MK-ELM	Proposed scheme	MK-ELM	Proposed scheme	MK-ELM	Proposed scheme
DoS	98.04	99.00	0.49	0.001	1.96	1.00	99.51	99.90
Probes	95.67	97.20	0.47	0.001	4.33	2.80	99.53	99.90
R2L	76.12	96.80	0.11	0.015	23.88	3.20	99.89	99.98
U2R	50.00	90.04	0.00	0.003	50.00	9.96	100.00	99.99

Table 12: Comparison performance of machine learning models using the CICIDS2017 benchmark dataset.

Classifier	ML models results				Hybrid PTE-BO ML model results			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
XGBoost	99.82	99.82	99.82	99.80	99.82	99.86	99.82	99.83
Ensemble stacking	99.82	99.91	99.82	99.85	99.82	99.91	99.82	99.85
Random forest	99.77	99.77	99.77	99.75	99.82	99.82	99.82	99.80
Decision tree	99.77	99.77	99.77	99.75	99.82	99.91	99.82	99.85
Extra tree	99.82	99.82	99.82	99.80	99.82	99.80	99.82	99.80

The validation of the result can also be confirmed by comparing the previous works as stated theoretically and graphically. The multilayer perception artificial neural network (MLPANN) effectively detects and classifies multiple attacks using public datasets, including UNSW-NB, WSN-DS, and NSL-KDD, as a benchmark for performance evaluation. By combining a tree based on the Parzen estimation (PTE) with hyperparameter and Bayesian optimization (BO) techniques, we are able to better classify the machine learning models for the proposed scheme on the benchmark dataset as shown in Table 12. Every single machine learning task uses hyperparameters to fine-tune the aforementioned parameters and get optimal results. Hyperparameter optimization (HPO) accomplishes both of these goals with less manual labor and better results from machine learning [77].

The MLPANN technique also achieves better detection accuracy of 99.62% using the WSN-DS benchmark dataset. The proposed scheme is effective for the localization and detection of different classes of attacks, approving that the proposed system has optimal average detection for multiple suspicious nodes. The novelty of this work is that it is effective in the detection and localization of various attacks. The proposed scheme is innovative for its ability to scale in both security and performance for optimal area coverage in wireless sensor networks with a hierarchical architecture and both heterogeneous and homogeneous sensor nodes.

6. Conclusion and Remarks

In this work, we proposed a multilayer perceptron artificial neural network (MLPANN) for detecting and localizing multiple attacks in WSNs. The proposed scheme achieved an average detection accuracy of 100%, 99.65%, 98.95%, and 99.83% for the various malicious nodes using UNSW-NB, WSN-DS, NSL-KDD, and CICIDS2018 benchmark datasets, respectively. The optimized localization approach is more effective and performs more significantly by 20% than the distance vector hop technique, with average localization accuracy of 99.12% using 160 beacon nodes. The validation of the proposed method is confirmed with the previous studies using the ANN classification technique using Python, IBM SPSS, and WEKA toolboxes for data processing and MATLAB R2021a for network planning and simulation. The datasets are used to evaluate the proposed system for detecting and localization accuracy of different attacks. The effectiveness of the proposed scheme is assessed using detection rate, ROC, false-positive rate, a lifetime of the network, residual energy, and the area under the curve metrics. The beacon, sensor, and malicious nodes were used hierarchically to simulate the target field. It is recommended to enhance further the detection and localization of accuracy of malicious nodes using different approaches in WSNs. We will extend this work with various attack classes and methods. The results show that performance and security of the proposed scheme are applicable for scalable and large network coverage in wireless sensor networks with heterogeneous and homogenous sensors for ensuring quality of services and availably. The proposed scheme will be examined in the future using other network planning and tools with different public datasets as benchmarks for detecting and localization attacks in WSNs.

Data Availability

The underlying dataset used to generate the results presented in this article is available upon request to the corresponding author.

Conflicts of Interest

The authors declare that there is no competing interest in this work.

References

[1] S. Messous and H. Liouane, "Online sequential DV-hop localization algorithm for wireless sensor networks," *Mobile Information Systems*, vol. 2020, Article ID 8195309, 14 pages, 2020.

[2] S. Dong, X. G. Zhang, and W. G. Zhou, "A security localization algorithm based on DV-hop against Sybil attack in wireless sensor networks," *Journal of Electrical Engineering and Technology*, vol. 15, no. 2, pp. 919–926, 2020.

[3] G. Farjamnia, Y. Gasimov, and C. Kazimov, "An improved DV-hop for detecting wormhole attacks in wireless sensor networks," *Journal of Communication Engineering*, vol. 9, no. 1, pp. 1–24, 2020.

[4] S. T. Patel and N. H. Ministry, "A Review: Sybil Attack Detection Techniques in WSN," in *4th International Conference on Electronics and Communication Systems (ICECs)*, pp. 4–8.

[5] J. Jiang, G. Han, C. Zhu, Y. Dong, and N. Zhang, "Secure localization in wireless sensor networks: a survey (invited paper)," *The Journal of Communication*, vol. 6, no. 6, pp. 460–470, 2011.

[6] B. Madagouda and R. Sumathi, "Artificial neural network approach using mobile agent for localization in wireless sensor networks," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 1137–1144, 2021.

[7] M. Bernas and B. Płaczek, "Fully connected neural networks ensemble with signal strength clustering for indoor localization in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 11, no. 12, Article ID 403242, 2015.

[8] D. Zhou, W. Liu, W. Zhou, and S. Dong, "Research on network traffic identification based on multi layer perceptron," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 12, no. 1, pp. 201–208, 2014.

[9] M. A. Alsheikh, S. Lin, D. Niyato, and H. P. Tan, "Machine learning in wireless sensor networks: algorithms, strategies, and applications," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 4, pp. 1996–2018, 2014.

[10] M. Wazid and A. K. Das, "An efficient hybrid anomaly detection scheme using K-means clustering for wireless sensor networks," *Wireless Personal Communications*, vol. 90, no. 4, pp. 1971–2000, 2016.

[11] M. Abdan and S. A. H. Seno, "Machine learning methods for intrusive detection of wormhole attack in mobile ad hoc network (MANET)," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 2375702, 12 pages, 2022.

[12] B. Ahmad, W. Jian, Z. A. Ali, S. Tanvir, and M. S. A. Khan, "Hybrid anomaly detection by using clustering for wireless sensor network," *Wireless Personal Communications*, vol. 106, no. 4, pp. 1841–1853, 2019.

[13] G. Farjamnia, Y. Gasimov, and C. Kazimov, "Review of the techniques against the wormhole attacks on wireless sensor networks," *Wireless Personal Communications*, vol. 105, no. 4, pp. 1561–1584, 2019.

[14] R. Singh, J. Singh, and R. Singh, "Fuzzy based advanced hybrid intrusion detection system to detect malicious nodes in wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2017, Article ID 3548607, 14 pages, 2017.

[15] M. Mahajan, K. T. V. Reddy, and M. Rajput, "Design and simulation of a blacklisting technique for detection of hello flood attack on LEACH protocol," *Procedia Computer Science*, vol. 79, pp. 675–682, 2016.

[16] U. Jain and M. Hussain, "Securing wireless sensors in military applications through resilient authentication mechanism," *Procedia Computer Science*, vol. 171, no. 2019, pp. 719–728, 2020.

[17] S. Z. Wang, Y. Li, and W. Cheng, "Distributed classification of localization attacks in sensor networks using exchange-based feature extraction and classifier," *Journal of Sensors*, vol. 2016, Article ID 8672305, 18 pages, 2016.

[18] L. Chelouah, F. Semchedine, and L. Bouallouche-Medjkoune, "Localization protocols for mobile wireless sensor networks: a survey," *Computers and Electrical Engineering*, vol. 71, pp. 733–751, 2018.

[19] A. Hadir, K. Zine-Dine, M. Bakhouya, and J. El Kafi, "An improved DV-hop localization algorithm for wireless sensor networks," in *2018 13th IEEE conference on industrial electronics and applications (ICIEA)*, pp. 330–334, Wuhan, China, June 2014.

[20] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, "WSN-DS: a dataset for intrusion detection systems in wireless sensor networks," *Journal of Sensors*, vol. 2016, Article ID 4731953, 16 pages, 2016.

[21] S. T. Patel and N. H. Mistry, "A review: Sybil attack detection techniques in WSN," in *2017 4th International conference on electronics and communication systems (ICECS)*, vol. 17, pp. 184–188, Coimbatore, India, February 2017.

[22] O. Cheikhrouhou and A. Koubaa, "BlockLoc: secure localization in the internet of things using blockchain," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 629–634, Tangier, Morocco, June 2019.

[23] F. Y. Yavuz, D. Ünal, and E. Gül, "Deep learning for detection of routing attacks in the internet of things," *International Journal of Computational Intelligence Systems*, vol. 12, no. 1, pp. 39–58, 2018.

[24] V. Sujatha and E. A. M. Anita, "FEM-hybrid machine learning approach for the detection of sybil attacks in the wireless sensor networks," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 7, pp. 1171–1179, 2019.

[25] X. Qi, X. Liu, and L. Liu, "A combined localization algorithm for wireless sensor networks," *Mathematical Problems in Engineering*, vol. 2018, Article ID 4648109, 10 pages, 2018.

[26] P. Li, X. Yu, H. Xu, J. Qian, L. Dong, and H. Nie, "Research on secure localization model based on trust valuation in wireless sensor networks," *Security and Communication Networks*, vol. 2017, Article ID 6102780, 12 pages, 2017.

[27] L. Song, L. Zhao, and J. Ye, "DV-hop node location algorithm based on GSO in wireless sensor networks," *Journal of Sensors*, vol. 2019, Article ID 2986954, 9 pages, 2019.

[28] M. Saud Khan and N. M. Khan, "Low complexity signed response based Sybil attack detection mechanism in wireless sensor networks," *Journal of Sensors*, vol. 2016, Article ID 9783072, 9 pages, 2016.

[29] A. Payal, C. S. Rai, and B. V. R. Reddy, "Artificial neural networks for developing localization framework in wireless sensor networks," in *2014 international conference on data mining and intelligent computing (ICDMIC)*, p. 5, Delhi, India, September 2014.

[30] R. Dela Cruz, "Artificial neural network-based localization in wireless sensor networks artificial neural network-based localization in wireless sensor networks," p. 14, 2019.

[31] J. Jiang, G. Han, H. Wang, and M. Guizani, "A survey on location privacy protection in wireless sensor networks," *Journal of*

Network and Computer Applications, vol. 125, pp. 93–114, 2019.

[32] R. Rastogi, S. Srivastava, M. S. Tarun, V. Manshahia, and N. Kumar, "A withdrawn: a hybrid optimization approach using PSO and ant colony in wireless sensor network," *Materials Today: Proceedings*, vol. 2021, 2021.

[33] Z. Han, L. Lin, Z. Wang et al., "CNN-based attack defense for device-free localization," *Mobile Information Systems*, vol. 2022, Article ID 2323293, 7 pages, 2022.

[34] J. Chen, S. H. Sackey, J. H. Anajemba, X. Zhang, and Y. He, "Energy-efficient clustering and localization technique using genetic algorithm in wireless sensor networks," *Complexity*, vol. 2021, Article ID 5541449, 12 pages, 2021.

[35] O. Cheikhrouhou, G. M. Bhatti, and R. Alroobaea, "A hybrid DV-hop algorithm using RSSI for localization in large-scale wireless sensor networks," *Sensors*, vol. 18, no. 5, pp. 1–14, 2018.

[36] W. He, F. Lu, J. Chen, Y. Ruan, T. Lu, and Y. Zhang, "A kernel-based node localization in anisotropic wireless sensor network," *Scientific Programming*, vol. 2021, Article ID 9944358, 8 pages, 2021.

[37] B. K. Madagouda and R. Sumathi, "Analysis of localization using ANN models in wireless sensor networks," in *2019 IEEE Pune Section International Conference (PuneCon)*, pp. 18–21, Pune, India, December 2019.

[38] R. Singh, J. Singh, and R. Singh, "Detection of Sybil Nodes in wireless sensor networks," *Journal of Advanced Computer Science & Technology*, vol. 10, no. 3, pp. 185–202, 2017.

[39] N. Moustafa, G. Creech, and J. Slay, "Anomaly detection system using beta mixture models and outlier detection," in *Progress in Computing, Analytics and Networking. Advances in Intelligent Systems and Computing*, P. Pattnaik, S. Rautaray, H. Das, and J. Nayak, Eds., vol. 710, Springer, Singapore, 2018.

[40] "Cybersecurity_attacks|Kaggle," 2021, https://www.kaggle.com/iamranjann/cybersecurity-attacks.

[41] Z. Zoghi, G. Serpen, and C. Science, "UNSW-NB15 computer security dataset: analysis through visualization".

[42] N. Moustafa, G. Creech, and J. Slay, "Flow aggregator module for analysing network traffic," in *Progress in Computing, Analytics and Networking. Advances in Intelligent Systems and Computing*, P. Pattnaik, S. Rautaray, H. Das, and J. Nayak, Eds., vol. 710, Springer, Singapore, 2018.

[43] "The UNSW-NB15 data set description," 2021, https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/.

[44] N. Moustafa and J. Slay, "A network forensic scheme using correntropy-variation for attack detection," *IFIP Advances in Information and Communication Technology*, vol. 532, pp. 225–239, 2018.

[45] N. Moustafa, G. Creech, and J. Slay, "Big data analytics for intrusion detection system: statistical decision-making using finite Dirichlet mixture models," in *Data Analytics and Decision Support for Cybersecurity*, pp. 127–156, Springer, 2017.

[46] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1–3, pp. 18–31, 2016.

[47] D. Praveen Kumar, T. Amgoth, and C. S. R. Annavarapu, "Machine learning algorithms for wireless sensor networks: a survey," *Information Fusion*, vol. 49, pp. 1–25, 2019.

[48] C. R. Panigrahi, "Advanced computing and intelligent engineering," *Proceedings of ICACIE*, vol. 1, 2018.

[49] M. N. A. Shaon and K. Ferens, "A computationally intelligent approach to the detection of wormhole attacks in wireless sensor networks," *Advances in Science, Technology and Engineering Systems*, vol. 2, no. 3, pp. 302–320, 2017.

[50] M. Mittal, R. P. de Prado, Y. Kawai, S. Nakajima, and J. E. Muñoz-Expósito, "Machine learning techniques for energy efficiency and anomaly detection in hybrid wireless sensor networks," *Energies*, vol. 14, no. 11, 2021.

[51] R. Ahmad, R. Wazirali, and T. Abu-Ain, "Machine learning for wireless sensor networks security: an overview of challenges and issues," *Sensors*, vol. 22, no. 13, p. 4730, 2022.

[52] "Artificial neural network - an overview | ScienceDirect Topics," 2021, https://www.sciencedirect.com/topics/engineering/artificial-neural-network.

[53] V. Hassija, S. Batra, V. Chamola et al., "A blockchain and deep neural networks-based secure framework for enhanced crop protection," *Ad Hoc Networks*, vol. 119, article 102537, 2021.

[54] K. Ncibi, T. Sadraoui, M. Faycel, and A. Djenina, "A multilayer perceptron artificial neural networks based a preprocessing and hybrid optimization task for data mining and classification," *International Journal of Economics, Finance and Management Sciences*, vol. 5, no. 1, pp. 12–21, 2017.

[55] M. Al-Imran and S. H. Ripon, "Network intrusion detection: an analytical assessment using deep learning and state-of-the-art machine learning models," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, pp. 1–20, 2021.

[56] D. Kothona, I. P. Panapakidis, and G. C. Christoforidis, "A novel hybrid ensemble LSTM-FFNN forecasting model for very short-term and short-term PV generation forecasting," *IET Renewable Power Generation*, vol. 16, no. 1, pp. 3–18, 2022.

[57] A. Hadir, Y. Regragui, and N. M. Garcia, "Accurate range-free localization algorithms based on PSO for wireless sensor networks," *IEEE Access*, vol. 9, pp. 149906–149924, 2021.

[58] H. Chen, W. Lou, Z. Wang, J. Wu, Z. Wang, and A. Xi, "Securing DV-Hop localization against wormhole attacks in wireless sensor networks," *Pervasive and Mobile Computing*, vol. 16, pp. 22–35, 2015.

[59] S. Awan, N. Javaid, S. Ullah, A. U. Khan, A. M. Qamar, and J. G. Choi, "Blockchain based secure routing and trust management in wireless sensor networks," *Sensors*, vol. 22, no. 2, pp. 1–24, 2022.

[60] G. Kumar, M. K. Rai, H. J. Kim, and R. Saha, "A secure localization approach using mutual authentication and insider node validation in wireless sensor networks," *Mobile Information Systems*, vol. 2017, Article ID 3243570, 12 pages, 2017.

[61] M. Cheng, T. Qin, and J. Yang, "Node localization algorithm based on modified Archimedes optimization algorithm in wireless sensor networks," *Journal of Sensors*, vol. 2022, Article ID 7026728, 18 pages, 2022.

[62] Z. Ansari, R. Ghazizadeh, and Z. Shokhmzan, "Gradient descent approach to secure localization for underwater wireless sensor networks," in *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, pp. 103–107, Shiraz, Iran, May 2016.

[63] "Classification: accuracy|Machine Learning Crash Course," 2021, https://developers.google.com/machine-learning/crash-course/classification/accuracy.

[64] D. Parikh and V. Menon, "Machine learning applied to cervical cancer data," *International Journal of Mathematics and Computer Science*, vol. 5, no. 1, pp. 53–64, 2019.

[65] "What are TP rate, FP rate, precision, recall, F measure, MCC, ROC area and PRC areas in the Weka tool?- Quora," 2020, https://www.quora.com/What-are-TP-rate-FP-rate-precision-recall-F-measure-MCC-ROC-area-and-PRC-areas-in-the-Weka-tool.

[66] "ROC analysis," 2021, https://www.ibm.com/support/knowledgecenter/SSLVMB_sub/statistics_mainhelp_ddita/spss/base/idh_roc.html.

[67] W. Zhang, D. Han, K. C. Li, and F. I. Massetto, "Wireless sensor network intrusion detection system based on MK-ELM," *Soft Computing*, vol. 24, no. 16, pp. 12361–12374, 2020.

[68] F. A. Khan, A. H. Farooqi, and A. Derhab, "A comprehensive security analysis of LEACH++ clustering protocol for wireless sensor networks," *The Journal of Supercomputing*, vol. 75, no. 4, pp. 2221–2242, 2019.

[69] L. Mohammadpour, T. C. Ling, C. S. Liew, and A. Aryanfar, "A mean convolutional layer for intrusion detection system," *Security and Communication Networks*, vol. 2020, Article ID 8891185, 16 pages, 2020.

[70] L. Xinlong and C. Zhibin, "DDoS attack detection by hybrid deep learning methodologies," *Security and Communication Networks*, vol. 2022, Article ID 7866096, 7 pages, 2022.

[71] B. Hasan, S. Alani, and M. A. Saad, "Secured node detection technique based on artificial neural network for wireless sensor network," *International Journal of Electrical & Computer Engineering*, vol. 11, no. 1, pp. 536–544, 2021.

[72] A. M. Pasikhani, J. A. Clark, and P. Gope, "Reinforcement-learning-based IDS for 6LoWPAN," in *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 1049–1060, Shenyang, China, 2022.

[73] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting feature selection with machine learning classifiers for intrusion detection on power grids," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1104–1116, 2021.

[74] S. P. K. Gudla, S. K. Bhoi, S. R. Nayak, and A. Verma, "DI-ADS: a deep intelligent distributed denial of service attack detection scheme for fog-based IoT applications," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3747302, 17 pages, 2022.

[75] M. I. Alghamdi, "A hybrid model for intrusion detection in IoT applications," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 4553502, 9 pages, 2022.

[76] R. Khilar, K. Mariyappan, M. S. Christo et al., "Artificial intelligence-based security protocols to resist attacks in Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1440538, 10 pages, 2022.

[77] M. Feurer and F. Hutter, "Hyperparameter optimization," in *Automated Machine Learning. The Springer Series on Challenges in Machine Learning*, F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., Springer, Cham, Switzerland, 2019.